

(19) 대한민국특허청(KR)(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)  
(12) 공개특허공보(A)

(51) Int. Cl. <sup>7</sup>	(11) 공개번호	특2002-0030693
H04Q 7/38(51) Int. Cl. <sup>7</sup>	(43) 공개일자	2002년04월25
H04Q 7/38	일(11) 공개번호	특2002-0030693
	(43) 공개일자	2002년04월25일
(21) 출원번호(21) 출원	10-2001-001009010-2001-0010090	
번호		
(22) 출원일자(22) 출원	2001년02월27일2001년02월27일	
일자		
(54) 음성 통역 서비스 방법 및 음성 통역 서버(30) 우선권주장		
2000-321921 2000년10월17일 일본(JP)		

요약

본 발명의 목적은, 통화용 전화기, 특히, 보급율이 높은 모바일 인터넷 액세스 서비스 이용이 가능한 전화를 사용하여 통역 정확도가 높은 음성 통역을 누구나 간단히 이용할 수 있는 서비스를 제공하는 것에 있다. 상기 목적을 달성하기 위해, 본 발명에서는 사용자가 전화를 이용하여 서버상에서 가동하는 자동 통역 서버에 음성을 입력한 후, 송신하면, 번역된 음성을 다시 유저의 전화에 회신하도록 한다. 그 때에, 최초에 유저가 모바일 인터넷 액세스 서비스 이용이 가능한 전화로부터 모바일 인터넷 액세스 서비스 패킷 망을 통해, 모바일 인터넷 액세스 서비스 게이트웨이 서버에 접속하면, 상기 자동 통역 서버는, 유저가 통역 가능한 언어 종류, 유저가 통역 가능한 문례(예시 문장 혹은 문장에) 장면, 유저가 입력 가능한 문례를, 상기 유저가 갖고 있는 전화의 표시 화면에 표시하도록 한다.

대표도

도1

색인어

모바일 인터넷, 자동 음성 통역, 문례, 인증 서버, 통역 서버

영세서

도면의 간단한 설명

- 도 1은 본 발명의 제1 실시예를 도시하는 일영(日英) 자동 통역 서비스 시스템 구성도.
- 도 2는 메모리(15)의 데이터 구조의 예.
- 도 3은 전화 단말(1)의 예.
- 도 4는 전화 단말(1)의 디스플레이(101)에의 서비스 메뉴 표시예.
- 도 5는 전화 단말(1)의 디스플레이(101)에의 통역 가능한 언어 종류 표시예.
- 도 6은 전화 단말(1)의 디스플레이(101)에의 통역 가능한 장면 표시예.
- 도 7은 전화 단말(1)의 디스플레이(101)에의 통역 가능한 문례(文例) 표시예.
- 도 8은 전화 단말(1)의 디스플레이(101)에의 인식 결과 표시예.
- 도 9는 언어 변환 테이블(10)의 데이터 구조의 예.
- 도 10은 전화 단말(1)의 디스플레이(101)에의 통역 결과 표시예.
- 도 11은 요금 부과 테이블(33)의 데이터 구조의 예.
- 도 12는 언어 종류 테이블(22)의 데이터 구조의 예.
- 도 13은 장면 테이블(23)의 데이터 구조의 예.
- 도 14는 문례 테이블(24)의 데이터 구조의 예.
- 도 15는 문장 사전(20)의 데이터 구조의 예.
- 도 16은 커맨드 사전(21)의 데이터 구조의 예.
- 도 17은 본 발명의 제2 실시예를 도시하는 일영 자동 통역 서비스 시스템 구성도.

도 18은 제1 실시예의 동작을 도시하는 흐름도.

도 19는 제1 실시예의 동작을 도시하는 흐름도.

(도면의 주요 부분에 대한 부호의 설명)

- 1 : 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말
- 2 : 모바일 인터넷 액세스 서비스 패킷 망
- 3 : 모바일 인터넷 액세스 서비스 게이트웨이 서버
- 4 : 전화 회선 보드
- 5 : 음성 입력 수단
- 6 : 음성 인식 수단
- 7 : 언어 번역 수단
- 8 : 단어 사전
- 9 : 문법 테이블
- 10 : 언어 변환 테이블
- 11 : 음성 생성 수단
- 12 : 소편 세트(음성 분절 세트)
- 13 : 음성 출력 수단
- 14 : 중앙 연산 장치
- 15 : 메모리
- 16 : 언어 종류 표시 수단
- 17 : 장면 표시 수단
- 18 : 문례(예시 문장) 표시 수단
- 19 : 인식 후보 표시 수단
- 31 : 인증 서버
- 32 : 요금 부과 서버
- 33 : 요금 부과 테이블

#### 발명의 상세한 설명

##### 발명의 목적

##### 발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 유저가 발성한 음성을 다른 언어로 번역하여 음성으로서 출력하는 자동 통역 서비스에 관한 것이다. 주로 모바일 인터넷 액세스 서비스 이용이 가능한 휴대 전화를 이용한 서비스를 대상으로 하지만, 통상의 유선 전화를 이용하거나, 퍼스널 컴퓨터, 휴대 정보 단말 등을 이용하여 상기 동작을 행하는 음성 통역 서비스도 포함한다.

종래, 음성 통역기라는 것이 있었다. 예를 들면, 특개평 11-125539에는, 내장 마이크로폰으로부터 입력되어 음성을 다른 언어로 번역하고, 번역 결과를 마이크로폰과는 반대 면에 설치된 스피커로부터 음성 출력하는, 소형이면서 또한 한 손 조작이 가능한 음성 통역 장치가 개시되어 있다. 그러나, 이러한 음성 통역 장치는 전용기(專用機)이기 때문에, 다른 용도로는 이용할 수 없고, 해외 여행 등의 실제 이용 시에는, 결국 유저의 소지품을 늘려 버리는 것이 된다. 한편, 예를 들면, 특개평9-65424에는, 음성 인식 서버와 무선 휴대 단말을 이용한 음성 번역 시스템이 개시되어 있다. 그러나, 이 음성 번역 시스템에서는, 임의 문장을 입력 대상으로 하고 있기 때문에, 높은 통역 정밀도가 얻어지기 어렵고, 실제 운용이 곤란하다는 문제가 있었다.

##### 발명이 이루고자하는 기술적 과제

본 발명은 상기한 점을 감안하여 이루어진 것으로, 통화용 전화기, 특히, 보급율이 높은 모바일 인터넷 액세스 서비스 이용이 가능한 전화를 이용하여, 정밀도가 높은 음성 통역을 누구나 간단히 이용할 수 있는 서비스를 제공하는 데에 그 목적이 있다.

##### 발명의 구성 및 작용

본 발명에서는, 유저가 전화를 이용하여 서버상에서 가동되는 자동 통역 서버에 음성을 송신하면, 번역된 음성이, 다시 유저의 전화에 회신되도록 한다. 그때에, 최초로 유저가 모바일 인터넷 액세스 서비스

이용이 가능한 전화로부터 모바일 인터넷 액세스 서비스 패킷망을 통하여, 모바일 인터넷 액세스 서비스 게이트웨이 서버에 접속하면, 상기 자동 통역 서버가, 사용자가 통역 가능한 언어 종류를, 사용자가 갖고 있는 전화의 표시 화면에 메뉴로서 표시하여, 유저로 하여금 상기 언어 종류 메뉴로부터 통역 희망 언어를 선택할 수 있도록 한다. 또한, 유저가 통역 가능한 문례 장면을 상기 유저의 전화의 표시 화면에 표시하여, 유저가 상기 장면 메뉴로부터 통역 가능 문례 장면을 선택할 수 있도록 한다. 또한, 유저가 입력 가능한 문례를 상기 유저의 전화의 표시 화면에 표시하여, 유저가 상기 입력 가능한 문례를 보면서 발성 입력을 할 수 있도록 한다. 또한, 상기 자동 통역 서버는, 상기 입력된 음성을 상기 표시한 범위의 문례 사전을 이용하여 인식하고, 역어(譯語)로 변환하여, 음성으로서 상기 전화 단말에 출력한다.

이와 같이, 새롭게 전용기를 구입하지 않고 일반 전화 단말을 이용하여, 상기 전화 단말에 표시한 정형적(定型的)인 문장을 입력 대상으로 하여, 자동 통역 서버로 통역 처리를 함으로써, 통역 정확도가 높은 통역 서비스를 용이하게 받을 수 있는 환경을 제공할 수가 있다.

#### (실시의 형태)

이하, 본 발명의 실시예를 도면을 이용하여 설명한다.

도 1은 본 발명의 제1 실시예를 도시하는 자동 통역 서비스 시스템의 구성도이다. 본 발명은, 음성 통역기에 관한 것으로, 본 실시예에 있어서의 서버는, 퍼스널 컴퓨터, 워크스테이션 등, 중앙 연산 장치, 메모리를 구비하며, 동도와 같은 구성을 할 수 있는 것이면 어느 것이나 좋고, 플랫폼의 종류가 본 발명의 적용 범위를 한정하는 것은 아니다. 또한, 본 실시예는 영어를 일본어로 통역하는 자동 통역 서비스이지만, 이것은 일례이고, 일영, 중일, 일중 등, 언어의 종류는 제한되지 않는다.

동 도면에 있어서, 참조 부호 1은 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말, 2는 모바일 인터넷 액세스 서비스 패킷망, 3은 모바일 인터넷 액세스 서비스 게이트웨이 서버, 4는 전화 회선 보드, 5는 음성 입력 수단, 6은 음성 인식 수단, 7은 언어 번역 수단, 8은 단어 사전, 9는 문법 테이블, 10은 언어 변환 테이블, 11은 음성 생성 수단, 12는 소편(素片) 세트(speech segments set: 음성 분절 세트), 13은 음성 출력 수단, 14는 중앙 연산 장치, 15는 메모리, 16은 언어 종류 표시 수단, 17은 장면 표시 수단, 18은 문례 표시 수단, 19는 인식 후보 표시 수단, 20은 문장 사전, 21은 커맨드 사전, 22는 언어 종류 테이블, 23은 장면 테이블, 24는 문례 테이블, 31은 인증 서버, 32는 요금 부과 서버, 33은 요금 부과 테이블이다. 메모리(15)의 데이터 구조를 도 2에 도시한다. 또한, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)의 개관의 예를 도 3에 도시한다. 또, 여기서는, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말로서, 무선, 통화 음성 및 데이터를 동일한 프로토콜로 다룰 수 있는 전화 단말을 가정하여 설명한다.

도 1, 도 3에 있어서, 우선, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)의 전원(104)을 넣어, 모바일 인터넷 접속용 버튼(102)을 누르면, 모바일 인터넷 액세스 서비스 패킷망(2)을 통해, 센터의 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에 접속되고, 인증 서버(31)에서 모바일 인터넷 액세스 서비스 이용 등록 유저인 것이 확인된다. 이 시점에서, 유저 ID가 요금 부과 서버(32)에 보내어진다. 당연하지만, 유저 ID는 전화 단말(1)의 ID와 링크하고 있는 것이 많지만, 전화 단말(1)과는 무관하게 유저 ID를 패스워드와 조합하여 입력시키는 방법도 있다.

요금 부과 서버(32)는 요금 부과 테이블(33)을 갖는다. 요금 부과 테이블(33)의 데이터 구조를 도 11에 도시한다. 인증 서버(31)로부터 보내어진 ID는, 요금 부과 테이블(33)의 유저 ID 컬럼(411)의 각 셀과 대조되어, 일치한 셀의 요금 컬럼(412)이 제로 리셋된다. 예를 들면, 상기 유저의 유저 ID가 '1236'이라고 하면, 요금 부과 테이블(33)의 유저 ID 컬럼(411)의 셀 중 '1236'인 셀(403)과 일치하기 때문에, 요금 컬럼(412)의 403에 대응하는 셀이 '0'이 된다.

모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에는, 본 실시예의 자동 통역 서버와 같은 각종 서비스 서버가 전용선으로 접속되어 있거나, 일반 인터넷 망이 접속되어 있다. 또, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에의 접속에는, 상기와 같은 모바일 인터넷 접속용 버튼(102)을 누르는 방법 외에, 어떤 소정의 전화번호를 숫자 버튼으로 입력하는 방법 등을 이용하여도 좋다.

모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)이, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에 접속되어, 인증 서버(31)에서 모바일 인터넷 액세스 서비스 이용 등록 유저인 것이 확인되면, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)는, 전화 단말(1)의 디스플레이(101)에, 도 4와 같이 서비스의 메뉴를 표시한다. 상기 서비스 메뉴는, 도 4와 같이, 흑백 반전에 의해 1번째 항목의 휘도가 다른 항목과는 다른 상태에서 초기 표시된다. 이것은, 제1번째의 항목이 선택되어 있는 것을 나타내지만, 물론, 그 표시는 흑백 반전에 한하지 않고, 다른 항목과 구별이 되어질 수 있는 방법이면 좋다.

유저는, 상기 서비스 메뉴를 보면서 전화 단말(1)상의 상하 커서 이동의 기능이 할당된 소정의 버튼을 눌러, 흑백 반전을 제3번째의 항목 '자동 통역'까지 이동시키고, 또한 전화 단말(1)상의 상기 소정의 버튼과는 다른 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 메뉴 항목 '자동 통역'이 확정되면, 전화 단말(1)이, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)를 경유하여 자동 통역 서버(1000)에 접속된다.

다음에, 자동 통역 서버(1000)의 언어 종류 표시 수단(16)이 기동되고, 언어 종류 테이블(22)을 이용하여 전화 단말(1)상의 디스플레이(101)에, 도 5와 같이 통역 가능한 언어의 종류를 표시한다. 언어 종류 테이블(22)은 도 12와 같은 데이터 구조를 갖고 있고, 언어 종류 표시 수단(16)은 상기 테이블의 언어 종류(12)의 각 항목을, 상기 전화 단말(1)에 보내어 도 5와 같이 상기 전화 단말(1)상의 디스플레이(101)에 표시한다. 동 도면에서는, 흑백 반전에 의해 1번째의 항목의 휘도가 다른 항목과는 다른 상태에서 초기 표시되어 있다. 이것은, 제1번째의 항목이 선택되어 있는 것을 나타내지만, 물론, 그 표시는 흑백 반전에 한하지 않고, 다른 항목과 구별이 되어질 수 있는 것 같은 방법이면 좋다. 유저는, 상기 언어 종류 메뉴를 보면서 전화 단말(1)상의 상하 커서 이동의 기능을 할당한 소정의 버튼을 누



러, 흑백 반전을 예를 들면 제2번째의 항목 '일본어 → 영어'까지 이동시키고, 또한 전화 단말(1)상의 상기 소정의 버튼과는 다른 소정의 확정 기능이 할당된 버튼을 눌러 확정한다. 이때, 언어 종류 표시 수단(16)은 전화 단말(1)의 커서의 위치를 수취하여 메모리(15)상의 LANG209에 상기 위치 번호를 저장한다. 예를 들면, 일본어를 영어로 통역하고 싶은 경우에는, 「일본어 → 영어」가 2번째에 있으므로, 메모리(15)상의 LANG209에 '2'가 저장된다.

물론, 이 언어 종류 지정은, 전화 단말(1)상의 디스플레이(101)에 언어 종류를 표시하여, 유저에게 원하는 언어 종류를 상하 이동 버튼으로 선택시키는 방법 외에, 상기 전화 단말(1)상의 디스플레이(101)를 터치 패널 디스플레이로 구성하고 손가락 등의 접촉에 의해 원하는 언어 종류를 지정하는 방법을 이용하도록 해도 좋다.

또한, 미리 어느 소정의 전화번호를 언어 종류마다 할당해 놓고, 유저가 원하는 언어 종류에 대응하는 전화번호를 상기 전화 단말(1)상의 숫자 버튼으로 입력하는 방법 등을 이용하여도 좋다.

언어 종류가 확정되면, 다음에, 자동 통역 서버(1000)의 장면 표시 수단(17)이 기동되고, 장면 테이블(23)을 이용하여, 전화 단말(1)상의 디스플레이(101)에, 도 6과 같이 통역 가능한 장면을 표시한다. 여기서 말하는 장면이란, 「공항」 「호텔」 「레스토랑」 등, 본 발명의 통역 서비스를 이용하는 장면의 것이다. 장면 테이블(23)은, 도 13과 같은 데이터 구조를 갖고 있고, 장면 표시 수단(17)은 상기 테이블의 장면(912)의 각 항목을, 상기 전화 단말(1)에 보내어 도 6과 같이 상기 전화 단말(1)상의 디스플레이(101)에 표시한다. 도 13에서는, 흑백 반전에 의해 1 번째의 항목의 휘도가 다른 항목과는 다른 상태에서 초기 표시되어 있다. 이것은, 제1번째의 항목이 선택되어 있는 것을 나타내지만, 물론, 그 표시는 흑백 반전에 한하지 않고, 다른 항목과 구별짓는 방법이면 좋다.

유저는, 상기 장면 메뉴를 보면서 전화 단말(1)상의 상하 커서 이동의 기능이 할당된 소정의 버튼을 눌러, 흑백 반전을 예를 들면 제3번째의 항목 '레스토랑'까지 이동시키고, 또한 전화 단말(1)상의 상기 소정의 버튼과는 다른 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 이때, 장면 표시 수단(17)은 전화 단말(1)의 커서의 위치를 수신하여 메모리(15)상의 SCENE210에 상태당 위치 번호를 저장한다. 예를 들면, 레스토랑의 장면으로 통역하고 싶은 경우에는, 「레스토랑」이 3번째에 있으므로, 메모리(15)상의 SCENE210에 '3'이 저장된다.

물론, 이 장면 지정은, 전화 단말(1)상의 디스플레이(101)에 장면을 표시하여, 유저에게 원하는 장면을 상하 이동 버튼으로 선택시키는 방법 외에, 상기 전화 단말(1)상의 디스플레이(101)를 터치 패널 디스플레이로 구성하고 손가락 등의 접촉에 의해 원하는 장면을 지정하는 방법을 이용하도록 해도 좋다. 또한, 미리 있는 소정의 전화번호를 장면마다 할당해 놓고, 유저가 원하는 장면에 대응하는 전화번호를 상기 전화 단말(1)상의 숫자 버튼으로 입력하는 방법 등을 이용하여도 좋다.

장면이 확정되면, 다음에, 자동 통역 서버(1000)의 문례 표시 수단(18)이 기동되고, 문례 테이블(24)을 이용하여 전화 단말(1)상의 디스플레이(101)에, 도 7과 같이 통역 가능한 문례를 표시함과 동시에, 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된다. 계속해서, 음성 입력 수단(5)은, 시스템을 음성 입력 가능한 상태로 한다. 그 후, 유저는, 상기 문례를 보면서, 예를 들면 전번의 예이면, 일본어로 레스토랑의 장면에서의 통역시키고 싶은 문장을 전화 단말(1)의 송화기의 마이크(104)를 향해서 발성 입력한다. 여기서는, 예로서, 레스토랑의 장면에서 「물을 원합니다」라는 문장을 영어로 통역하고 싶은 경우를 생각하는 것으로 한다.

문례 테이블(24)은 도 14와 같은 데이터 구조를 갖고 있고, 문례 표시 수단(18)은, 문례 테이블(24)의 장면 번호(511)의 각 항목 중, 메모리(15)상의 SCENE210에 저장되어 있는 값의 항목의 문례(513)를 문례 번호(512)의 '1'로부터 순차적으로, 상기 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 표시한다. 전번의 예에서는, 메모리(15)상의 SCENE210에 '3'이 저장되어 있기 때문에, 도 14의 문례 테이블(24)의 장면 번호(511)의 값이 '3'인 항목(501, 502, 503, 504)의 문례(513)의 내용「안녕하십니까」 「고맙습니다」 「[ ]는 어디입니까」 「[ ]를 원합니다」를 상기 전화 단말(1)에 보내어 도 7과 같이 상기 전화 단말(1)상의 디스플레이(101)에 M 개씩 순차 표시한다. M은 디스플레이(101)의 크기에 따라서 설정하는 상수이지만, 여기서는 M = 4으로 한다.

유저는, 상기 문례의 속에 「[ ]를 원합니다」라는 문례가 있기 때문에, 이 문례를 기초로 「물을 원합니다」로 발성 입력한다. 이 발성 입력에 있어서, 전화 단말(1)상의 발성 입력의 신호의 의미를 가지게 한 소정의 버튼을 눌러 자동 통역 서버(1000)의 음성 입력 수단(5)을 음성 입력 가능한 상태로 하고 나서 발성해도 좋고, 특히 버튼 등의 트리거를 이용하지 않고 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된 후 언제나 음성 입력 가능한 상태로 해도 좋다. 또, 표시하는 문례는, 상술한 바와 같이 슬롯 [ ]을 함유한 문장 혹은 문법이라도 좋고, 혹은 문장 그 자체라도 좋다. 여기서, 슬롯이란, 단어, 구 등을 그 장소에 설정할 수 있다고 하는 말하자면 박스이다. 예를 들면, 「[ ]를 원합니다」에서는, 슬롯 [ ]에 「물」 「커피」 「찬 물」 등을 설정할 수 있다.

이와 같이, 문례를 표시하여 문형을 어느 정도 규정함으로써, 음성 인식의 정확도를 현저히 향상시킬 수 있다. 또한, 문례를 표시함으로써 유저의 사용성을 향상시킬 수 있다.

또, 상기 표시 문례는, 상하 커서 이동의 기능을 할당한 소정의 버튼을 누름으로써, 차례 차례로 스크롤하고 인식 가능한 문례를 전부 순차 표시시킬 수 있다. 다만, 이때에, 상기 전화 단말(1)상의 디스플레이(101)에 표시시키고 있는 최초의 문례의 문례 번호(512)의 값, 및, 표시시키고 있는 최후의 문례의 문례 번호(512)의 값을 메모리(15)상의 BSENTENCE211, ESENTENCE212에 각각 저장한다. 도 7의 예에서는, BSENTENCE211에는 '1', ESENTENCE212에는 '4'가 저장된다.

계속해서, 음성 입력 수단(5)은, 전화 회선 보드(4)로서 아날로그/디지털 변환된 입력 음성을 메모리(15)상의 WAVE201에 저장한다. 전화 회선 보드(4)에서의 아날로그/디지털 변환의 샘플링률은 8 kHz, 11 kHz, 16 kHz 등, 유저가 적절하게 정할 수 있다.

또, 이 시점에서, 상기 유저가, 입력한 음성을 취소하고, 재 입력하고 싶다고 하면, 전화 단말(1)상의

취소 기능이 할당된 소정의 버튼을 누름으로써, 음성을 재 입력할 수 있는 상태로 할 수 있게 된다. 전화 단말(1)상의 취소 기능이 할당된 소정의 버튼은, 누르면, 발성 입력의 신호의 의미를 가진 소정의 버튼을 눌렀을 때 혹은 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된 직후와 동일한 초기 상태로 리셋되어, 음성 입력 가능한 상태가 되도록 한다.

다음에, 음성 인식 수단(6)이 기동된다. 음성 인식 수단(6)은, 메모리(15)상 WAVE201에 저장된 음성 데이터를 읽어들이, 특징 벡터의 시계열로 변환하고, 각 문장(sentence) 음성의 상기 특징 벡터 시계열을 갖는 문장 사전(20)을 이용하여 대조를 행하고, 상기 음성 데이터를 인식하여 후보를 출력한다. 특징 벡터 시계열에의 변환 방법이나, 대조 방법을 포함해서 음성을 인식하는 방법으로서, 예를 들면 L. Rabiner & B. -H. Juang저, 古井貞熙 監譯「음성 인식의 기초(하)」(NTT 어드반스 테크놀로지, 1995) PAGE 245~P304 기재의 방법 등을 이용할 수가 있다. 물론, 다른 음성 인식방법을 이용하여도 좋고, 음성 인식 방법이 본 발명을 한정하는 것은 아니다. 문장 사전(20)의 데이터 구조를 도 15에 도시한다. 음성 인식 수단(6)은, 메모리(15)상 WAVE201에 저장된 음성 데이터를 읽어들이, 문장 사전(20)의 문례 번호(611)의 값이 메모리(15)상의 BSENTENCE211, ESENTENCE212에 각각 저장되어 있는 값의 범위내인 항목의 것 전체의 특징 벡터 시계열(614)의 값을 이용하여 음성 인식을 행한다. 전번의 예에서는, BSENTENCE211에는 '1', ESENTENCE212에는 '4'가 저장되어 있기때문에, 문장 사전(20)의 문례 번호611의 값이 '1'에서 '4'까지의 각 항목의 특징 벡터 시계열(614)의 값을 이용하여 음성 인식을 행한다. 그 결과, 상기 음성은, 매칭 정도가 높은 순서로, 예를 들면, 문자열 '메뉴를 원합니다', '물을 원합니다', '커피를 원합니다', '스폰을 원합니다'의 문례 번호 및 문장 번호로 변환된다. 그 때문에, 이들 후보의 문례 번호611, 문장 번호612, 및 문장613을 매칭도가 높은 순서로, 메모리(15)상의 RECOGNUM(1), RECOGSNUM(1), RECOGS(1), RECOGNUM(2), RECOGS(2), ..., RECOGNUM(N), RECOGSNUM(N), RECOGS(N)205에 저장하여 간다. 여기서, N은, 문장 사전(20)의 문례 번호611의 값이 메모리(15)상의 BSENTENCE211, ESENTENCE212에 각각 저장되어 있는 값의 범위내인 항목의 전체 수이다. 다음에, 인식 후보 표시 수단(19)이 기동되어, RECOGS(1), RECOGS(2), ..., RECOGS(M) 205의 내용을, 도 8과 같이, 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순서대로 표시한다. 이때, 메모리(15)상의 ICNT204에 '1'을 저장하고, 전화 단말(1) 상의 디스플레이(101)에 표시되어 있는 RECOGS(ICNT)의 내용이 흑백 반전되도록 한다. M은 디스플레이(101)의 크기에 따라서 설정하는 상수이지만, 여기서는 M = 4로 한다. 또한, 메모리(15)상의 INDEX215에 '0'을 저장한다.

유저는, 제1 후보의 표시가 자신이 발성한 내용과 일치하고 있는지 혹은 근접하다면, 그대로 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 제1 후보 표시가 정확하지 않은 경우, 유저는, 상하 커서 이동의 기능이 할당된 소정의 버튼을 누름으로써, 전화 단말(1)상의 디스플레이(101)상의 정확한 문자열이 표시되어 있는 부분에까지 흑백 반전 표시 영역을 아래쪽으로 이동시킨다. 즉, 하 이동 버튼을 누를 때마다, 메모리(15)상의 ICNT 204의 값을 인크리먼트하고, 전화 단말(1)상의 디스플레이(101)상에 동일하게 메모리(15)상의 RECOGS(ICNT)의 내용이 표시되어 있는 부분만을 흑백 반전 표시하도록 한다. 여기서, ICNT 204의 값이 M을 넘은 경우에는, 메모리(15)상의 INDEX215의 값에 'M'을 가산하고, 다음 M 개의 후보 RECOGS(INDEX+1), RECOGS(INDEX+2), ..., RECOGS(INDEX+M)를 메모리(15)상의 가운데부터 판독하며, 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차 표시한다. 이때, 메모리(15)상의 ICNT204에 '1'을 저장하고, 전화 단말(1)상의 디스플레이(101)에 표시되어 있는 RECOGS(INDEX+1), RECOGS(INDEX+2), ..., RECOGS(INDEX+M)중의 ICNT 번째의 표시가 흑백 반전되도록 한다. 이하, 마찬가지로 하여, 다음의 M 개의 후보를, 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차적으로 표시해 간다. 또, 도 15에 도시한 음성 인식에 사용하는 문장 사전(20)의 구조는 일례이고, 예를 들면, 문법과 단어 사전을 조합하는 방법도 있다. 그 경우에도, 메모리(15)상의 SCENE210에 저장된 장면 번호에 대응시켜, 장면별 문법이나 사전을 이용하도록 한다.

또한, 상 이동 버튼으로 흑백 반전 표시 영역을 위쪽으로 이동시키도록 하고, 백하여 표시 내용을 보이도록 해도 좋다. 즉, 상 이동 버튼을 누를 때마다, 메모리(15)상의 ICNT 204의 값을 디크리먼트하여, 전화 단말(1)상의 디스플레이(101)상의 RECOGS(INDEX+1), RECOGS(INDEX+2), ..., RECOGS(INDEX+M)중의 ICNT 번째의 표시부분만을 흑백 반전 표시하도록 한다.

이 후보의 선택은, 상기 전화 단말(1)상의 상하 이동 버튼을 이용하는 방법에 한하지 않고, 상기 전화 단말(1)상의 디스플레이(101)를 터치 패널 디스플레이로 구성하고 손가락 등의 접촉에 의해 후보를 지정하는 방법을 이용하도록 해도 좋다.

유저가, 흑백 반전되어 있는 표시가 자신이 발성한 내용과 일치하고 있거나 혹은 가까운 것으로 하여 확정 기능이 할당된 소정의 버튼을 눌러 확정하면, 메모리(15)상의 RECOGNUM(INDEX+ICNT), RECOGSNUM(INDEX+ICNT), RECOGS(INDEX+ICNT)의 값을, 동일하게 메모리(15)상의 PCAND213, SCAND214, JAPANESE203에 각각 저장한다.

도 8의 예에서는, '물을 원합니다'가, 제2행째에 표시되어 있기 때문에, 하 이동 버튼을 누름으로써 제2행째까지, 흑백 반전 영역을 이동시키고, 확정 버튼을 누른다. 그리고, INDEX는 '0' ICNT는 '2'가 되고, RECOGNUM(2), RECOGSNUM(2), RECOGS(2)의 값인 '4', '1', '물을 원합니다'가 메모리(15)상의 PCAND213, SCAND214, JAPANESE203에 각각 저장된다.

자신이 발성한 내용의 확인 방법에는, 이상 진술한 바와 같은 전화 단말(1)상의 디스플레이(101)상에 음성 인식 후보를 표시하는 방법 외에, 디스플레이(101)를 사용하지 않고서, 이하와 같은 방법을 이용할 수도 있다. 즉, 음성 인식 수단(6)이 매칭 정도가 높은 순으로 후보의 문례 번호(611), 문장 번호(612), 및 문(613)을 매칭 정도가 높은 순으로, 메모리(15)상의 RECOGNUM(1), RECOGSNUM(1), RECOGS(1), RECOGNUM(2), RECOGSNUM(2), RECOGS(2), ..., RECOGNUM(N), RECOGSNUM(N), RECOGS(N)205에 저장한 후, 음성 생성 수단(12)이 기동되어, 이때, 메모리(15)상의 JCNT208에 '1'을 저장하고, 메모리(15)상의 RECOGS(JCNT)를 먼저 판독하여, 상기 문자열을 합성 음성으로 변환하고, 상기 음성의 파형 데이터를 디지털/아날로그 변환에 의해 아날로그 데이터로 변환하여, 음성 출력 수단(13)을 통하여 음성으로서 전화 단말(1)에 보낸다. 문자열을 합성 음성으로 변환하기 위해서는, 예를 들면, J.Allen, M.S.Hunnicut and D.Klatt 등의 'From Text to Speech'(Cambridge University Press, 1987) P.16에서



P.150에 기재되어 있는 홀만트 합성 방식이나, 失頭의 「텍스트 음성 합성 기술의 최신 상황」(Interface, Dec., 1996) P.161내지 P.165에 기재되어 있는 파형 중첩 방식을 쓸 수 있다. 물론, 다른 텍스트 음성 합성 방식을 이용할 수도 있다. 또한, 인식 가능한 문례에 대응하는 음성을 미리 녹음하여, 메모리 등의 기억 장치에 저장하여 놓도록 해도 좋다.

유저는, 상기 전송된 음성을 전화 단말(1)상의 스피커(100)로부터 듣고, 상기 음성이 자신이 발성 입력한 내용과 합치하고 있으면, 전화 단말 상의 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 상기 음성이 자신이 발성 입력한 내용과 합치하지 않는 경우에는, 유저는 전화 단말(1)상의 다음 후보 제시의 기능이 할당된 소정의 버튼을 누르도록 한다. 자동 통역 서버(1000)중의 음성 생성 수단(12)은, 전화 단말(1)상의 다음 후보 제시 기능이 할당된 소정의 버튼이 눌리면, 메모리(15)상의 JCNT208을 인크리먼트하여, RECOGS(JCNT)를 판독하고, 상기 문자열을 합성 음성으로 변환하며, 상기 음성의 파형 데이터를 디지털/아날로그 변환에 의해 아날로그 데이터로 변환하여, 음성 출력 수단(13)을 통해 음성으로서 전화 단말(1)에 보낸다.

상술한 동작과 마찬가지로, 유저는, 상기 전송된 음성을 전화 단말(1)상의 스피커(100)로부터 듣고, 상기 음성이 자신이 발성 입력한 내용과 합치하고 있으면, 전화 단말(1)상의 확정 기능이 할당된 소정의 버튼을 눌러 확정하고, 상기 음성이 자신이 발성 입력한 내용과 합치하지 않으면, 유저는 전화 단말(1)상의 다음 후보 제시의 기능이 할당된 소정의 버튼을 눌러, 자신이 발성 입력한 내용에 합치하는 음성이 들릴 때까지, 이상의 동작을 반복한다.

확정 버튼을 누르면, 메모리(15)상의 RECOGS(ICNT)에 저장되어 있는 문자열을 동일하게 메모리(15)상의 JAPANESE 203에 저장한다. 이 확정 버튼을 누르는 대신에, 어느 소정의 단어 혹은 구(句) 혹은 문장을 유저가 발성 입력하는 방법을 이용해도 좋다. 즉, 유저는, 상기 전송된 음성을 전화 단말(1)상의 스피커(100)로부터 듣고, 상기 음성이 자신이 발성 입력한 내용과 합치하고 있으면, 합치하고 있다는 취지의 소정의 단어 혹은 구 혹은 문장을 확정 음성으로서 전화 단말(1)상의 마이크(104)로부터 발성 입력하고, 상기 음성이 자신이 발성 입력한 내용과 합치하지 않으면, 합치하고 있는 경우와는 다른 소정의 단어 혹은 구 혹은 문장을 비확정 음성으로서 전화 단말(1)상의 마이크(104)로부터 발성 입력하도록 하여, 자동 통역 서버(1000)중의 음성 인식 수단(6)이, 상술한 바와 같은 문장 입력과 마찬가지로 방법으로 음성을 인식하여, 전체 매칭 정도가 미리 설정한 임계치를 하회하는 경우 혹은 ICNT204의 값이 N을 넘은 경우, 커맨드 사전(21)과의 대조를 개시한다. 커맨드 사전(21)의 데이터 구조를 도 16에 도시한다. 입력 음성의 특징 벡터의 시계열을, 커맨드 사전(21)의 각 항목의 특징 벡터의 시계열과 대조하여, 매칭 정도가 가장 높은 후보의 커맨드 번호를 갖고 커맨드로 한다. 여기서는 예로서, 유저가 「확정」이라고 발성 입력한 경우를 상정하면, 먼저, 문장 사전(20)을 이용한 인식으로, 상기 음성의 특징 벡터와 각 항목의 특징 벡터의 대조로 매칭 정도가 미리 설정한 임계치를 하회하는 것이 되고, 커맨드 사전(21)의 각 항목의 특징 벡터의 대조가 행하여져서, 701의 항목이 인식 후보가 된다. 상기 항목의 커맨드 번호가 1이면, 그것은 「확정」을 나타내는 발성 입력이 된다.

확정 음성이면, 메모리(15)상의 RECOGS(ICNT)에 저장되어 있는 문자열을 동일하게 메모리(15)상의 JAPANESE203에 저장한다. 또한, 비확정 음성이면, 메모리(15)상의 JCNT208을 인크리먼트하고, RECOGS(JCNT)를 판독하여, 상기 문자열을 합성 음성으로 변환하며, 상기 음성의 파형 데이터를 디지털/아날로그 변환에 의해 아날로그 데이터로 변환하여, 음성 출력 수단(13)을 통해 음성으로서 전화 단말(1)에 보낸다고 하는 상술한 동작을 확정 음성이 될 때까지 반복하도록 한다.

다음에, 자동 통역 서버(1000) 중의 언어 번역 수단(7)이 기동된다. 언어 번역 수단(7)은, 언어 변환 테이블(10)을 이용하여 메모리상의 JAPANESE203에 저장되어 있는 문자열을 다른 언어로 번역 변환한다. 이하, 언어 번역 수단(7)의 동작에 관해서 설명한다. 언어 변환 테이블(10)의 데이터 구조를 도 9에 도시한다.

우선, 언어 번역 수단(7)은, 메모리(15)상의 PCAND213, SCAND214에 저장되어 있는 값을 언어 변환 테이블(10)의 문례 번호311 및 문장 번호312의 각 항목과 순차 대조하여, 일치한 항목의 역어312 중의 LANG209의 값의 컬럼의 내용을 메모리(15)의 RESULT206에 저장한다.

앞의 예에서는, PCAND213, SCAND214에 저장되어 있는 값은 각각 '4', '1'이므로, 303의 항목 '물을 원합니다'와 일치한다. 또한, LANG209의 값은 '2'이기 때문에, 대응하는 역어312의 'I'd like to have a glass of water'를 메모리(15)의 RESULT206에 저장한다. 역어로의 변환에 대해서는, 상술한 예와 같은 언어 변환 테이블을 이용하는 방법에만 한하지 않고, 예를 들면 특개평03-328585에 기재되어 있는 음례 데이터 베이스를 이용한 번역 방법이나, 특개평03-51022에 기재되어 있는 구문 해석에 의한 번역 방법 등을 사용하더라도 좋다.

다음에, 언어 번역 수단(7)은, 도 10에 도시한 바와 같이, 메모리(15)의 JAPANESE 203 및 RESULT 206에 각각 저장되어 있는 내용을 전화 단말(1)상의 디스플레이(101)에 표시한다. 동 도면의 표시 방법은 일례이고, 표시 방법이 본 발명의 범위를 한정하는 것은 아니다.

계속해서, 자동 통역 서버(1000) 중의 음성 생성 수단(12)이 기동된다. 음성 생성 수단(12)은 메모리(15)상의 ENGLISH 206에 저장되어 있는 문자열을 읽어들이, 상기 문자열을 합성 음성으로 변환하고, 메모리(15)상의 SYNWAVE 207에 파형 데이터를 저장한다. 문자열을 합성 음성으로 변환하기 위해서는, '예를 들면, J.Allen, M.S.Hunnicutt and D.Klatt 등의 'From Text to Speech' (Cambridge University Press, 1987) P.16 내지 P.150에 기재되어 있는 홀만트 합성 방식이나, 失頭의 「텍스트 음성 합성 기술의 최신 상황」(Interface, Dec., 1996) P.161 내지 P.165에 기재되어 있는 파형 중첩 방식을 쓸 수 있다. 물론, 다른 텍스트 음성 합성 방식을 사용할 수도 있다. 또한, 각 영역(英譯)에 대응하는 음성을 미리 녹음하여, 그대로 혹은 압축하여 메모리 등의 기억 장치에 저장하여 놓도록 해도 좋다.

계속해서, 음성 생성 수단(12)은, 메모리(15)상의 SYNWAVE 207에 저장된 역어 음성의 파형 데이터를 아날로그 데이터 혹은 패킷 데이터로 변환하고, 음성 출력 수단(13)을 통하여 음성으로서 전화 단말(1)에 보내어, 상기 전화 단말(1)이 갖는 메모리에 상기 보내어진 역어 음성을 저장한다. 다만, 음성 출력 수

단(13)으로부터 출력된 역어 음성을 상기 전화 단말(1)이 갖는 메모리에 저장하지 않고, 자동 통역 서버(1000)의 메모리(15)상에 저장하도록 해도 좋다.

이 시점에서, 요금 부과 테이블(33)의 유저 ID 컬럼(411)의 상기 인증 서버(31)로부터 전송된 ID 셀에 대응하는 요금 컬럼 412의 내용에 통역 1회분의 소정의 요금이 가해진다. 예를 들면 통역 일회당 50엔이라는 요금을 설정하고 있으면, 전번의 예에서는, 유저 ID가 '1236'이므로, 요금 부과 테이블(33)의 유저 ID 컬럼411의 셀 중 '1236'인 셀(403)에 대응하는 요금 컬럼412의 셀이 '50'이 된다. 이 요금은, 1통역 서비스당 금액으로 설정해도 좋고, 어느 정액 요금으로 몇번이나 통역 서비스를 이용할 수 있도록 설정해도 좋다. 서비스 종료 후, 가산된 요금은, 각 ID의 유저에 대하여 청구되어 지불이 행하여진다. 요금 중의 수수료 등으로서 소정의 금액이 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)의 소유자에게 지불되고, 남은 금액이 자동 통역 서버(1000)의 소유자에게 지불되도록 한다. 다만, 이것들의 요금의 분배 방법이나 분배처는 일례이고, 상황에 따라서 설정할 수가 있다.

여기서, 유저가 전화 단말(1)상의 역어 출력의 기능을 할당한 소정의 버튼을 누르면, 상기 전화 단말(1)상의 메모리에 저장되어 있는 역어 음성을 판독하여, 상기 전화 단말(1) 상의 스피커(100)로부터 상기 역어 음성이 출력되도록 한다. 또한 유저가 전화 단말(1)상의 역어 출력의 기능이 할당된 소정의 버튼을 누르면, 그 때마다 상기 전화 단말(1)상의 메모리에 저장되어 있는 역어 음성을 판독하여, 전화 단말(1)상의 스피커(100)로부터 상기 역어 음성이 출력되도록 한다.

상기 역어 음성을 출력하기 위한 방법에는, 전화 단말(1)상의 역어 출력의 기능이 할당된 소정의 버튼을 누르는 방법에 한하지 않고, 소정의 단어 혹은 구 혹은 문장을 유저가 발성 입력하는 방법을 이용하여도 좋다. 즉, 유저는, 상기 역어 음성의 출력을 지시하는 취지의 소정의 단어 혹은 구 혹은 문장을 출력 지시 음성으로서 전화 단말(1)상의 마이크(104)로부터 발성 입력하고, 자동 통역 서버(1000)중의 음성 인식 수단(6)이 상기 음성을 인식하여, 출력 지시 음성이면, 상기 전화 단말(1)상의 메모리에 저장되어 있는 역어 음성을 판독하고, 전화 단말(1)상의 스피커(100)로부터 상기 역어 음성이 출력되도록 한다. 다만, 음성 출력 수단(13)으로부터 출력된 역어 음성이 상기 전화 단말(1)이 갖는 메모리에 저장되어 있지 않은 경우에는, 자동 통역 서버(1000)의 메모리(15)상에 저장되어 있는 역어 음성을 상기 전화 단말(1)에 보내도록 한다.

또, 여기서 진술한 상기 전화 단말(1)상의 메모리에 저장되어 있는 역어 음성을 판독하여 전화 단말(1)상의 스피커(100)로부터 상기 역어 음성을 출력시키는 경우에는, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에 어떤 정보를 보내는 일은 없고, 따라서 요금 부과 서버(32)에 의해서 상기 유저에 요금 부과가 이루어지는 일은 없다. 이상의 동작의 흐름을 도 18, 도 19에 흐름도로 도시한다.

계속해서, 본 발명의 제2 실시예에 관해서 도면을 이용하여 설명한다. 도 17은 본 발명의 제2 실시예를 도시하는 자동 통역 서비스 시스템의 구성도이다. 제1 실시예와 같이, 본 실시예에 있어서의 서버는, 퍼스널 컴퓨터, 워크스테이션 등, 중앙 연산 장치, 메모리를 구비하고, 동 도면과 같은 구성을 할 수 있는 것이면 어느 것이나 좋고, 플랫폼의 종류가 본 발명의 적용 범위를 한정하는 것은 아니다. 또한, 본 실시예는 영어를 일본어로 통역하는 자동 통역 서비스이지만, 이것은 일례이고, 일영, 중일, 일중 등, 언어의 종류는 한정되지 않는다. 동 도면에 있어서, 참조부호 1은 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말, 2는 모바일 인터넷 액세스 서비스 패킷 망, 3은 모바일 인터넷 액세스 서비스 게이트웨이 서버, 4는 전화 회선 보드, 5는 음성 입력 수단, 6은 음성 인식 수단, 7은 언어 번역 수단, 8은 단어 사전, 9는 문법 테이블, 10은 언어 변환 테이블, 11은 음성 생성 수단, 12는 소편 세트, 13은 음성 출력 수단, 14는 중앙 연산 장치, 15는 메모리, 16은 언어 종류 표시 수단, 17은 장면 표시 수단, 18은 문례 표시 수단, 19는 인식 후보 표시 수단, 20은 문장 사전, 21은 커맨드 사전, 22는 언어 종류 테이블, 23은 장면 테이블, 24는 문례 테이블, 31은 인증 서버, 32는 요금 부과 서버, 33는 요금 부과 테이블, 34는 통화용 통신망, 1000은 자동 통역 서버, 1001은 모바일 인터넷 액세스 서비스 패킷 망에 접속되는 회선, 1002는 통화용 통신망에 접속되는 회선이다.

메모리(15)의 데이터 구조 및 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)의 개관의 예는 제1 실시예와 마찬가지로이다. 다만, 여기서는, 제1 실시예와는 달리, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말로서, 통화 음성 및 데이터를 동일한 프로토콜로 취급하지 않는 전화 단말을 가정하여 설명한다.

도 17, 도 3에 있어서, 먼저, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)의 전원(104)을 넣고, 모바일 인터넷 접속용 버튼(102)을 누르면, 모바일 인터넷 액세스 서비스 패킷망(2)을 통해, 센터의 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에 접속되어, 인증 서버(31)로 모바일 인터넷 액세스 서비스 이용 등록 유저인 것이 확인된다. 이하, 장면 확정까지의 동작은, 제1 실시예와 마찬가지로이다.

장면이 확정되면, 다음에, 자동 통역 서버(1000)의 문례 표시 수단(18)이 기동되고, 문례 테이블(24)을 이용하여 회선(1001)을 통해 전화 단말(1)상의 디스플레이(101)에, 도 7과 같이 통역 가능한 문례를 표시함과 동시에, 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된다. 계속해서, 음성 입력 수단(5)은, 시스템을 음성 입력 가능한 상태로 한다. 그 후, 유저는, 상기 문례를 보면서, 예를 들면 전번의 예이면, 일본어로 레스토랑의 장면에서의 통역시키고 싶은 문장을 전화 단말(1)의, 송화기의 마이크(104)를 향해서 발성 입력한다. 여기서는, 예로서, 레스토랑의 장면에서 「물을 원합니다」라는 문장을 영어로 통역하고 싶은 경우를 생각하는 것으로 한다. 문례 테이블(24)은 도 14와 같은 데이터 구조를 갖고 있고, 문례 표시 수단(18)은, 문례 테이블(24)의 장면 번호511의 각 항목 중, 메모리(15)상의 SCENE210에 저장되어 있는 값의 항목의 문례513를 문례 번호512의 '1'로부터 순차적으로, 상기 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 표시한다. 전번의 예에서는, 메모리(15)상의 SCENE210에 '3'이 저장되어 있기 때문에, 도 14의 문례 테이블(24)의 장면 번호511의 값이 '3'인 항목501, 502, 503, 504의 문례(513)의 내용 「안녕하십니까」 「고맙습니다」 「[ ]는 어디입니까」 「[ ]를 원합니다」를 상기 전화 단말(1)에 보내어 도 7과 같이 상기 전화 단말(1)상의 디스플레이(101)에 M 개씩 순차 표시한다. M은 디스플레이(101)의 크기에 따라서 설정하는 상수이지만, 여기서는 M = 4로



한다.

유저는, 상기 문례의 중에 「[ ]를 원합니다」 라는 문례가 있기 때문에, 이 문례를 기초로 「물을 원합니다」 라고 발성 입력한다. 이 발성 입력에 즈음하여, 전화 단말(1)상의 발성 입력의 신호의 의미를 가 지게한 소정의 버튼을 눌러 자동 통역 서버(1000)의 음성 입력 수단(5)이 음성 입력 가능한 상태가 되고 나서 발성해도 좋고, 특히 버튼 등의 트리거를 이용하지 않고 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된 후 언제나 음성 입력 가능한 상태이더라도 좋다. 또, 표시하는 문례는, 상술한 예와 같이 슬롯[ ]을 함유한 문장 혹은 문법이라도 좋고, 혹은 문장 그 자체라도 좋다.

제2 실시예에서는, 통화 음성 및 데이터를 동일한 프로토콜로 다룰 수 없는 전화 단말을 가정하고 있기 때문에, 발성 입력한 음성을 자동 통역 서버에 보내기 위해서는, 통화용 통신망(20)을 통할 필요가 있다. 따라서, 유저는, 상기 전화 단말(1)로부터, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)를 경유하고 회선(1001)을 통하여 자동 통역 서버(1000)에 접속한 것과는 다른 전화번호로 통화용 통신망(34)을 경유하고 회선(1002)을 통해 자동 통역 서버에 접속한다. 여기서, 유저가 다시 접속하는 방법 외에, 자동 통역 서버(1000)의 음성 입력 수단(5)이 상기 유저의 전화 단말(1)에 접속하도록 해도 좋다. 하여간, 유저가 발성한 음성은, 통화용 통신망(34)을 경유하고 회선(1002)을 통해 자동 통역 서버(1000)에 보내어진다.

이하, 음성 인식 수단(6)이 기동되어 상기 발성 입력된 음성이 인식되고, 결과가 메모리에 저장되는 동작까지는, 제1 실시예와 마찬가지로이다. 다음에, 인식 후보 표시 수단(19)이 기동되어, RECOGS (1), RECOGS (2), ..., RECOGS (M)205의 내용을, 도 8과 같이, 회선(1001)을 통해 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차적으로 표시한다. 이 때, 메모리(15)상의 ICNT204에 '1'를 저장하고, 전화 단말(1)상의 디스플레이(101)에 표시되어 있는 RECOGS (ICNT)의 내용이 흑백 반전되도록 한다. M은 디스플레이(101)의 크기에 따라서 설정하는 상수이지만, 여기서는 M = 4로 한다. 또한, 메모리(15)상의 INDEX215에 '0'을 저장한다.

유저는, 제1 후보의 표시가 자신이 발성한 내용과 일치하고 있는지 혹은 가까우면, 그대로 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 제1 후보의 표시가 정확하지 않은 경우, 유저는, 상하 커서 이동 기능이 할당된 소정의 버튼을 누름으로써, 전화 단말(1) 상의 디스플레이(101)상의 정확한 문자열이 표시되어 있는 부분에까지 흑백 반전 표시 영역을 아래쪽으로 이동시킨다. 즉, 하 이동 버튼을 누를 때마다, 메모리(15)상의 ICNT 204의 값을 인크리먼트하여, 회선(1001)을 통해 전화 단말(1)의 디스플레이(101)상의 동일하게 메모리(15)상의 RECOGS(ICNT)의 내용이 표시되어 있는 부분만을 흑백 반전 표시하도록 한다. 여기서, ICNT 204의 값이 M을 넘은 경우에는, 메모리(15)상의 INDEX215의 값에 'M'을 가산하고, 다음 M 개의 후보 RECOGS(INDEX+1), RECOGS(INDEX+2), ..., RECOGS(INDEX+M)를 메모리(15)로부터 판독하여, 회선(1001)을 통해 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차적으로 표시한다. 이때, 메모리(15)상의 ICNT204에 '1'을 저장하고, 전화 단말(1)상의 디스플레이(101)에 표시되어 있는 RECOGS(INDEX+1), RECOGS(INDEX+2), ..., RECOGS(INDEX+M)중의 ICNT 번째의 표시가 흑백 반전되도록 한다. 이하, 마찬가지로 하여, 다음의 M 개의 후보를, 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차적으로 표시하여 간다. 또, 도 15에 도시한 음성 인식에 사용하는 문장 사전(20)의 구조는 일례이고, 예를 들면, 문법과 단어 사전을 조합하는 방법도 있다. 그 경우도, 메모리(15)상의 SCENE210에 저장된 장면 번호에 대응시켜, 장면별 문법이나 사전을 이용하도록 한다.

자신이 발성한 내용의 확인 동작은, 제1 실시예와 마찬가지로이지만, 자동 통역 서버(1000) 중의 음성 생성 수단(12)은, 전화 단말(1)상의 다음 후보 제시 기능이 할당된 소정의 버튼이 눌리면, 메모리(15)상의 JCNT208를 인크리먼트하여, RECOGS(JCNT)를 판독하고, 상기 문자열을 합성 음성으로 변환하며, 상기 음성의 파형 데이터를 디지털/아날로그 변환에 의해 아날로그 데이터로 변환하여, 음성 출력 수단(13)을 통해 음성으로서 회선(1001)을 통해 전화 단말(1)에 보낸다.

확정 버튼을 누르면, 메모리(15)상의 RECOGS (ICNT)에 저장되어 있는 문자열을 동일하게 메모리(15)상의 JAPANESE203에 저장한다. 확정 버튼의 신호를 자동 통역 서버(1000)에 보내는 루트는, 회선(1001)을 경유해도 좋고, 회선(1002)을 경유해도 좋다. 이 확정 버튼을 누르는 대신에, 어느 소정의 단어 혹은 구 혹은 문장을 유저가 발성 입력하는 방법을 이용하여도 좋다. 즉, 유저는, 상기 전송된 음성을 전화 단말(1)상의 스피커 입력(100)으로부터 듣고, 상기 음성이 자신이 발성 입력한 내용과 합치하고 있으면, 합치하고 있다는 취지의 소정의 단어 혹은 구 혹은 문장을 확정 음성으로서 전화 단말(1)상의 마이크(104)로부터 발성 입력하고, 상기 음성이 자신이 발성 입력한 내용과 합치하지 않으면, 합치하고 있는 경우와는 다른 소정의 단어 혹은 구 혹은 문장을 비확정 음성으로하여 전화 단말(1)상의 마이크(104)로부터 발성 입력하도록 하며, 회선(1002)을 통하여 자동 통역 서버(1000)에 보내어지고, 상기 자동 통역 서버 중의 음성 인식 수단(6)이, 상술한 바와 같은 문장 입력과 마찬가지로 방법으로 음성을 인식하여, 전체 매칭 정도가 미리 설정한 임계치를 하회하는 경우 혹은 ICNT204의 값이 N을 넘은 경우, 커맨드 사전(21)과의 대조를 개시한다. 다음에, 자동 통역 서버(1000) 중의 언어 번역 수단(7)이 기동되지만, 상기 언어 번역 수단(7)의 번역 동작은, 제1 실시예와 마찬가지로이다.

다음에, 언어 번역 수단(7)은, 도 10에 도시한 바와 같이, 메모리(15)의 JAPANESE203 및 RESULT206에 각각 저장되어 있는 내용을 회선(1001)을 통하여 전화 단말(1)상의 디스플레이(101)에 표시한다. 동 도면의 표시 방법은 일례이고, 표시의 방법이 본 발명의 범위를 한정하는 것은 아니다.

계속해서, 자동 통역 서버(1000) 중의 음성 생성 수단(12)이 기동되지만, 상기 음성 생성 수단(12)이 음성을 생성하는 동작은 제1 실시예와 마찬가지로이다.

계속해서, 음성 생성 수단(12)은, 메모리(15)상의 SYNWAVE207에 저장된 역어 음성의 파형 데이터를 아날로그 데이터 혹은 패킷 음성으로 변환하고, 음성 출력 수단(13)을 통해 음성으로서 회선(1002)을 통해 전화 단말(1)에 보내며, 상기 전화 단말(1)이 갖는 메모리에 상기 전송된 역어 음성을 저장한다. 이하, 요점 부과의 방법 및 동작에 관해서도, 제1 실시예와 마찬가지로이다.



## 발명의 효과

본 발명에 따르면, 휴대 전화 단말을 이용한 정확도가 높은 자동 통역 서비스를 실현할 수 있기 때문에, 상대의 언어를 인식하지 않고 이국인끼리의 상호 커뮤니케이션을 용이하게 수행할 수 있다.

## (57) 청구의 범위

### 청구항 1

음성 통역 서비스 방법에 있어서,

통역 서버와 모바일 인터넷 액세스망으로 접속된 휴대 단말의 화면에 상기 통역 서버에 저장된 통역용 문례집(通譯用文例集)을 표시시키는 단계와,

상기 휴대 단말로부터 음성 통역 서비스에의 접속 지시를 받아 상기 휴대 단말과 상기 통역 서버와의 접속을 전화망 접속으로 전환하는 단계와,

상기 휴대 단말에 입력된 유저 음성을 전화망을 통하여 수신하는 단계와,

상기 유저 음성을 인식하는 단계와,

인식한 유저 음성을 다른 언어 표현으로 통역하는 단계와,

통역된 다른 언어 표현에 대응하는 음성을 전화망을 통하여 상기 휴대 단말에 출력하는 단계를 포함하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 2

제1항에 있어서,

상기 모바일 인터넷 액세스망을 통하여, 상기 유저가 통역 가능한 언어 종류를, 상기 휴대 단말의 표시 화면에 메뉴로서 표시시키는 단계와,

유저로부터 상기 언어 종류 메뉴로부터 통역 희망 언어의 선택을 받는 단계와,

인식한 유저 음성을 선택된 언어의 표현으로 통역하는 단계를

를 포함하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 3

제1항에 있어서,

상기 모바일 인터넷 액세스망을 통하여 통역 가능한 문례 분야를 상기 휴대 단말의 표시 화면에 표시시키는 단계와,

유저로부터 상기 분야 메뉴로부터 통역 가능 문례 장면의 선택을 받는 단계와,

인식한 유저 음성을 선택된 문례 장면에 기초한 언어 표현으로 통역하는 단계를

를 포함하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 4

제1항에 있어서, 유저 음성을 인식한 결과를 전화망을 통하여 유저의 휴대 단말에 음성 출력하는 단계를 포함하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 5

제4항에 있어서,

상기 인식한 결과를 휴대 단말로 수신한 유저로부터의 지시를 받은 후, 통역 음성을 전화망을 통하여 상기 휴대 단말에 출력하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 6

제5항에 있어서, 상기 지시는, 상기 휴대 단말의 버튼을 밀어 내림에 따라 행해지고, 상기 유저가 상기 버튼을 누를때마다 통역 음성을 출력하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 7

제5항에 있어서, 상기 지시는, 상기 유저에 의한 특정한 단어 혹은 문장 음성의 발성 입력에 의해 행해지고, 상기 지시에 의해 통역 음성을 출력하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 8

제7항에 있어서, 상기 유저에 의한 특정 단어 혹은 문장 음성을 발성하고, 입력을 받을 때 마다, 통역 음성을 출력하는 것을 특징으로 하는 통역 서비스 방법.

### 청구항 9

제1항에 있어서,

음성 인식에 이용하는 사전 혹은 문법 혹은 이들 양쪽을 상기 휴대 단말의 표시 화면에 표시된 문례에 관한 항목에 조립하여 음성 인식을 행하는 것을 특징으로 하는 음성 통역 서비스 방법.

#### 청구항 10

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 통역 음성을 상기 휴대 단말에 출력하는 횟수에 무관하게, 유저에게 요금을 부과하는 것을 특징으로 하는 음성 통역 서비스 방법.

#### 청구항 11

음성 통역 서버에 있어서,

전화망을 통하여 음성의 입력을 받는 음성 입력 수단과,

입력된 음성을 인식하여 소정의 기호열로 변환하는 음성 인식 수단과,

변환된 기호 열을 상기 입력된 음성의 언어와 다른 언어의 표현으로 변환하는 언어 변환 수단과,

변환된 언어 표현에 대응하는 음성을 출력하는 음성 출력 수단과,

통역용 문례를 저장하는 문례 테이블과,

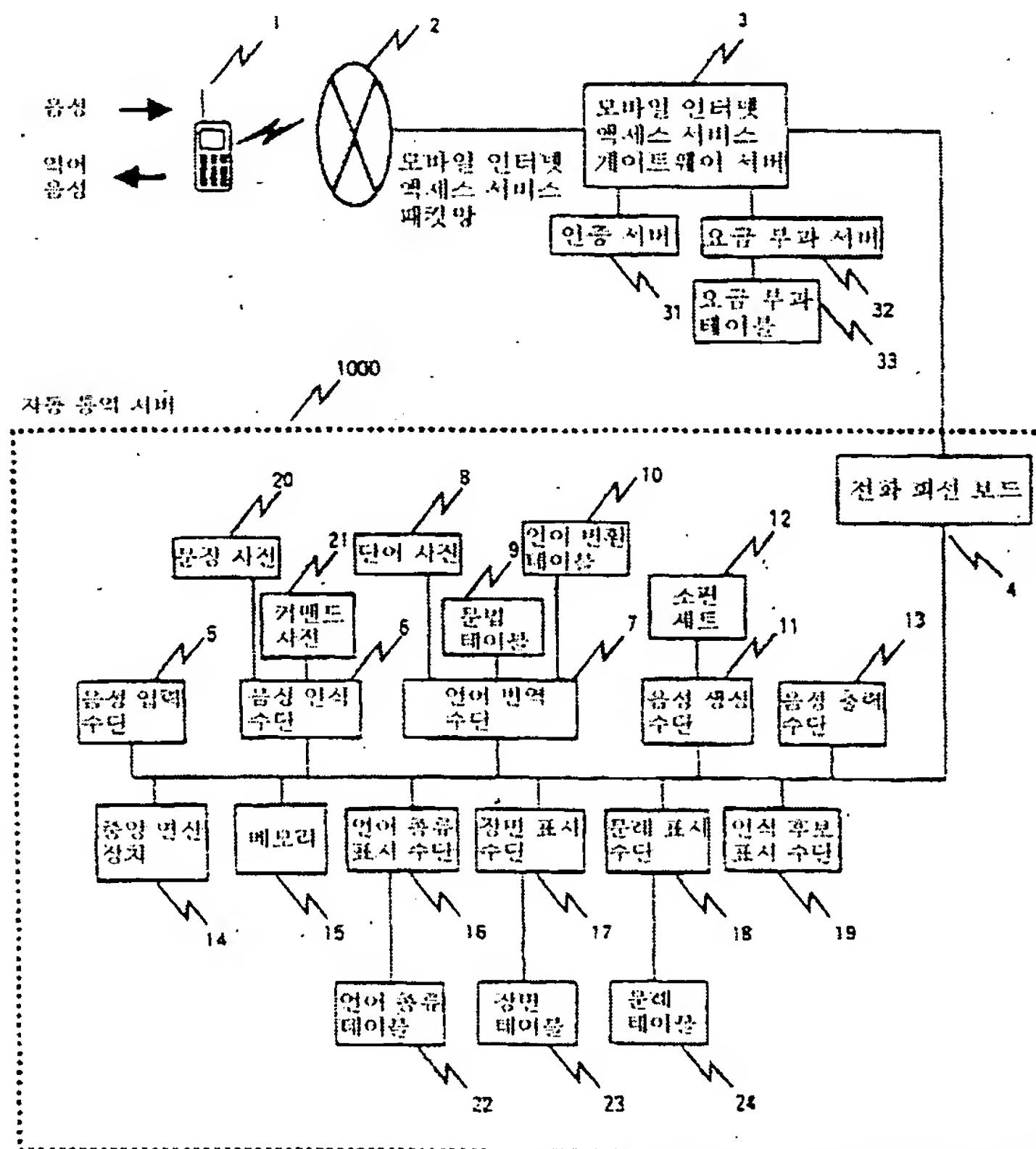
모바일 인터넷 액세스망으로 접속된 휴대 단말의 화면에 상기 통역용 문례를 표시시키는 문례 표시 수단을 포함하는 것을 특징으로 하는 음성 통역 서버.

#### 청구항 12

제11항에 있어서, 상기 음성 인식 수단은, 상기 문례 테이블을 참조하여 통역용 문례와, 동작을 지시하는 커맨드문을 구별하여 인식하는 것을 특징으로 하는 음성 통역 서버.

#### 도면

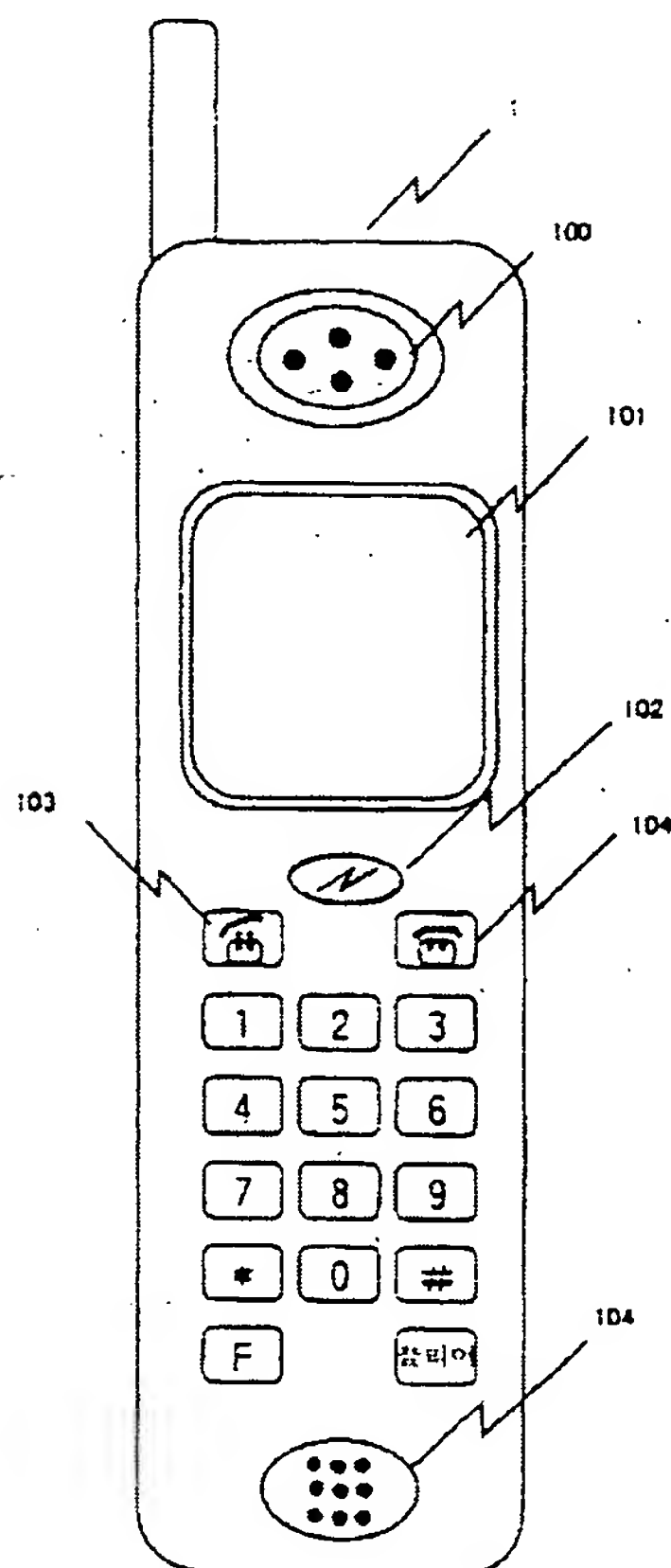
도면1



도면2

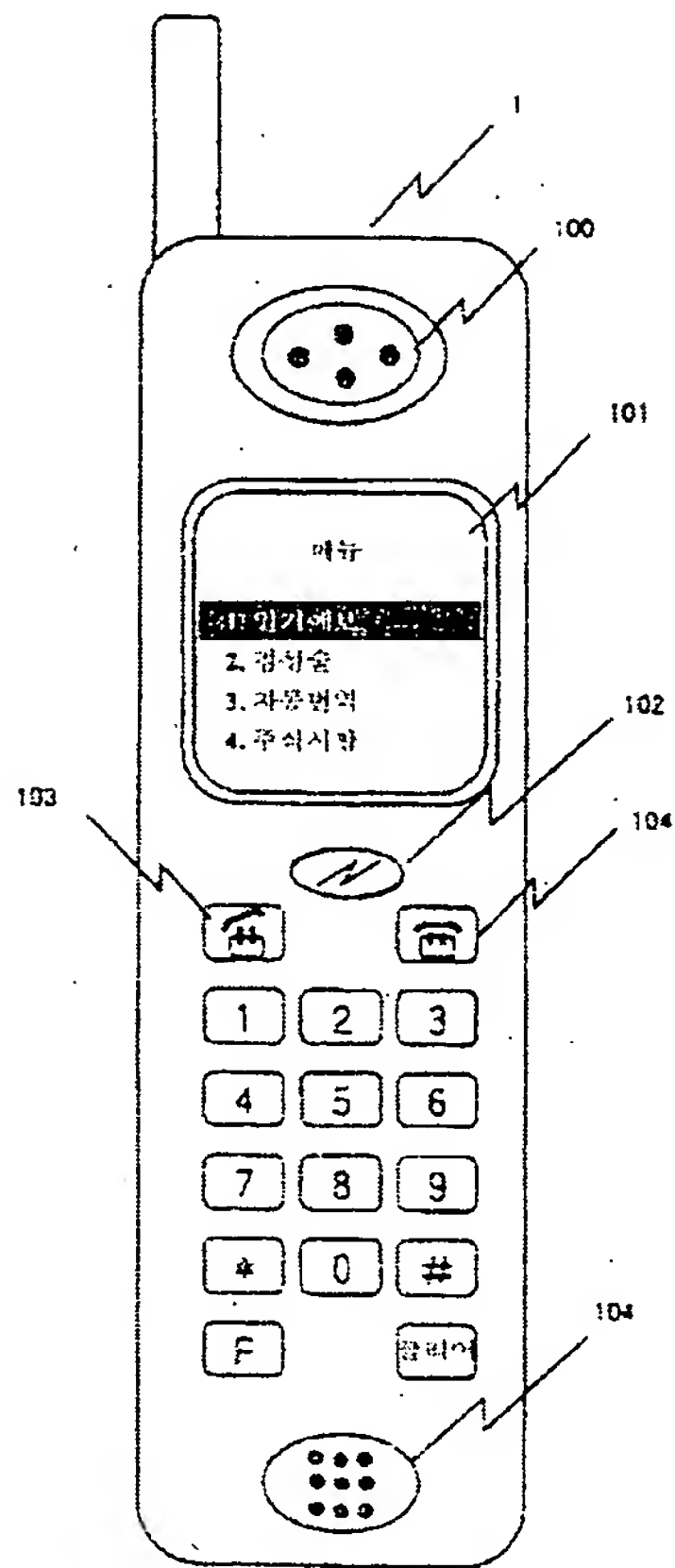
201	WAVE	
202	CSTRING	
203	JAPANESE	
204	JOINT	
205	RECOGS(1)	
	RECOGNUM(1)	
	RECOGSNUM(1)	
	:	
	RECOGS(N)	
	RECOGNUM(N)	
	RECOGSNUM(N)	
206	RESULT	
207	SYNWAIVE	
208	JOINT	
209	LANG	
210	SCENE	
211	BSENTENCE	
212	CSENTENCE	
213	PCAND	
214	SCAND	
215	INDEX	
	:	

도면3

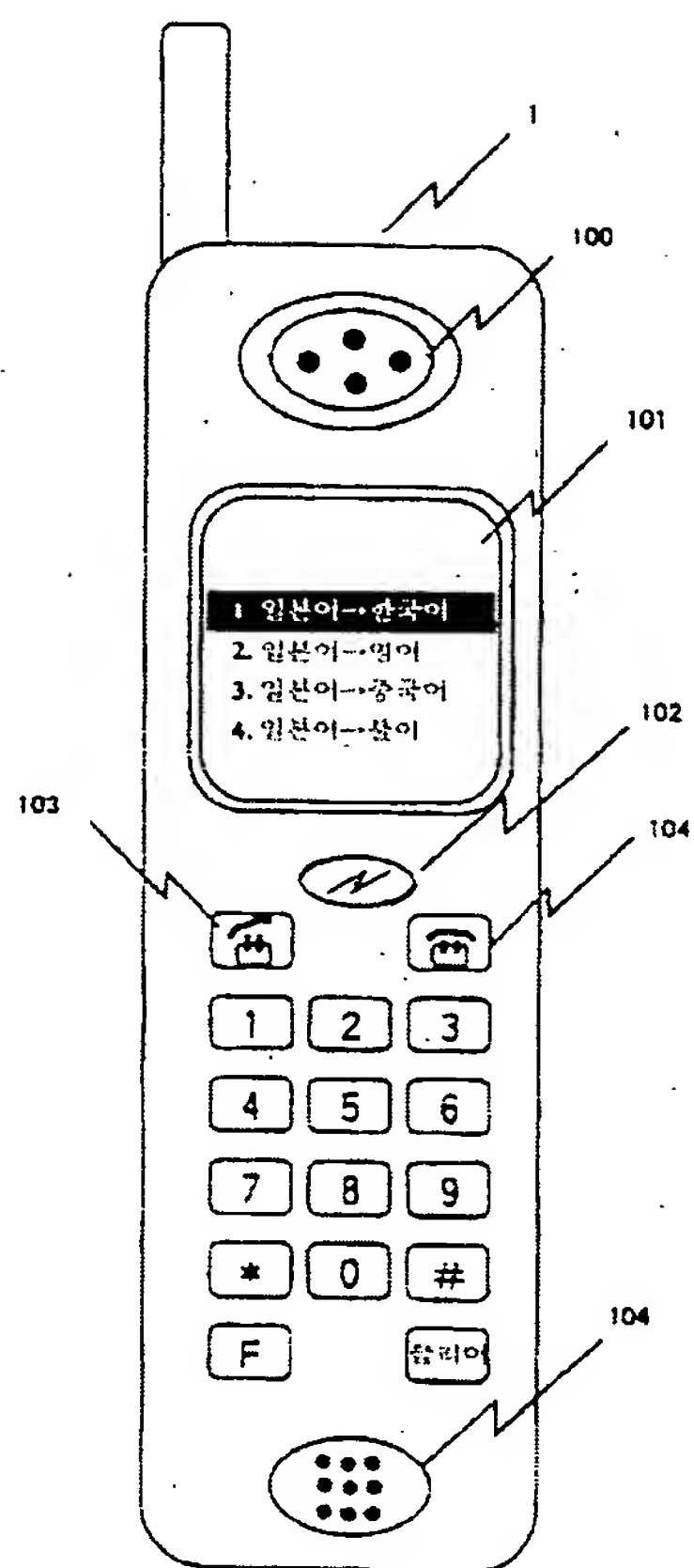




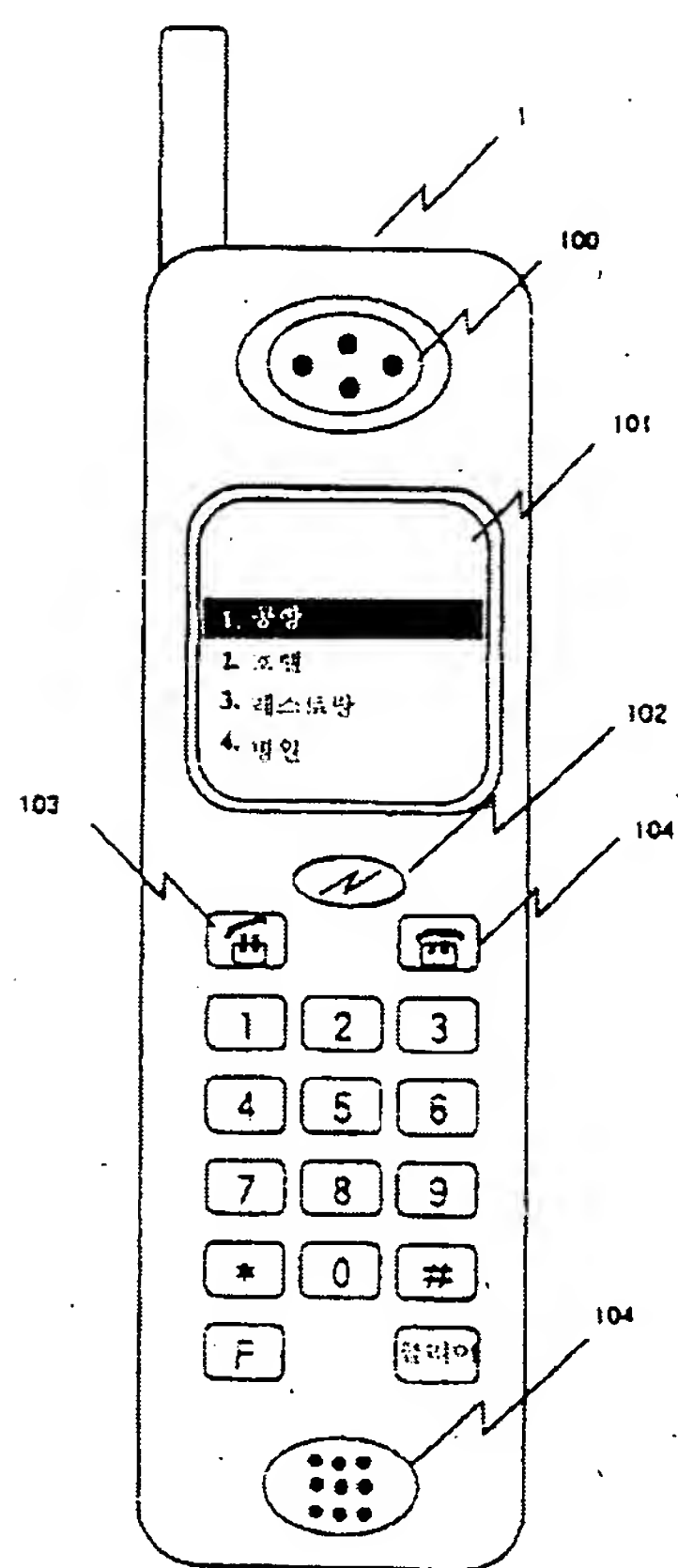
도면4



도면5

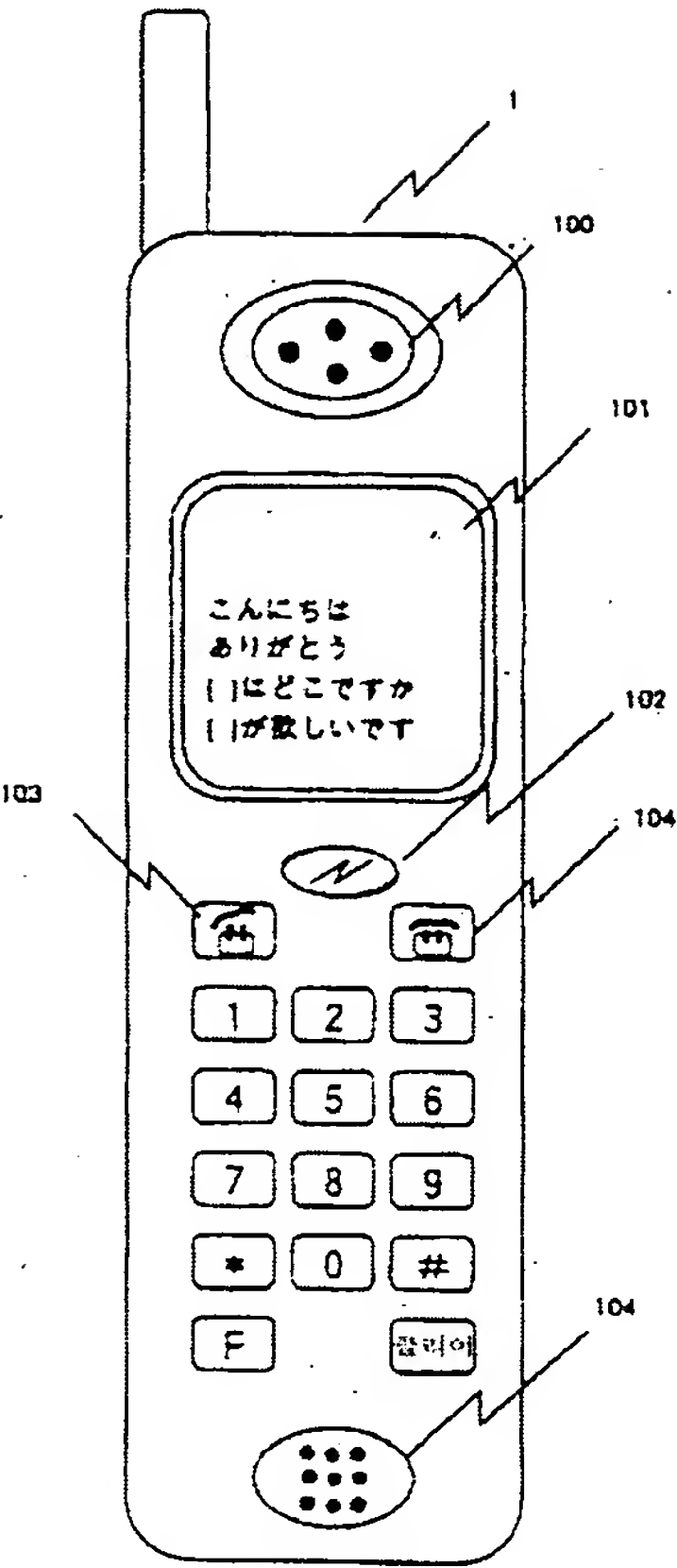


도면6

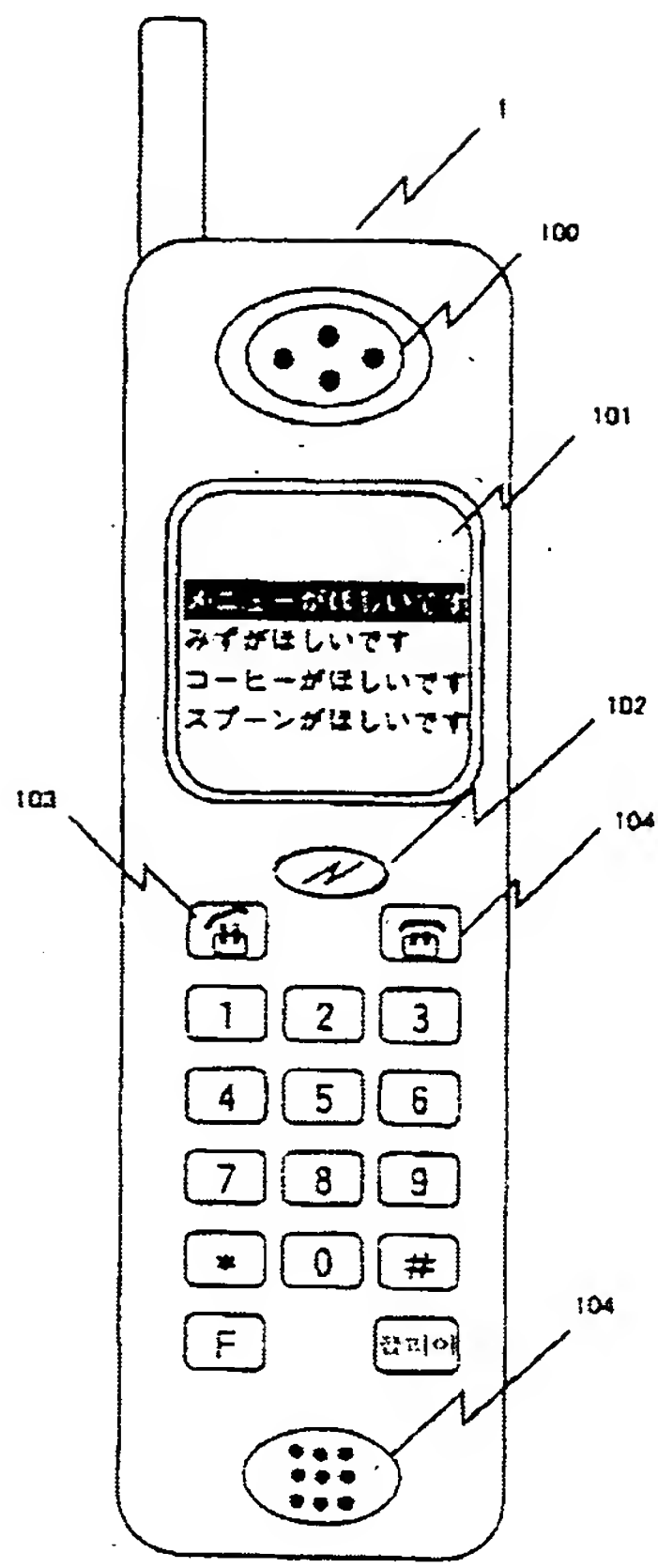




도면7



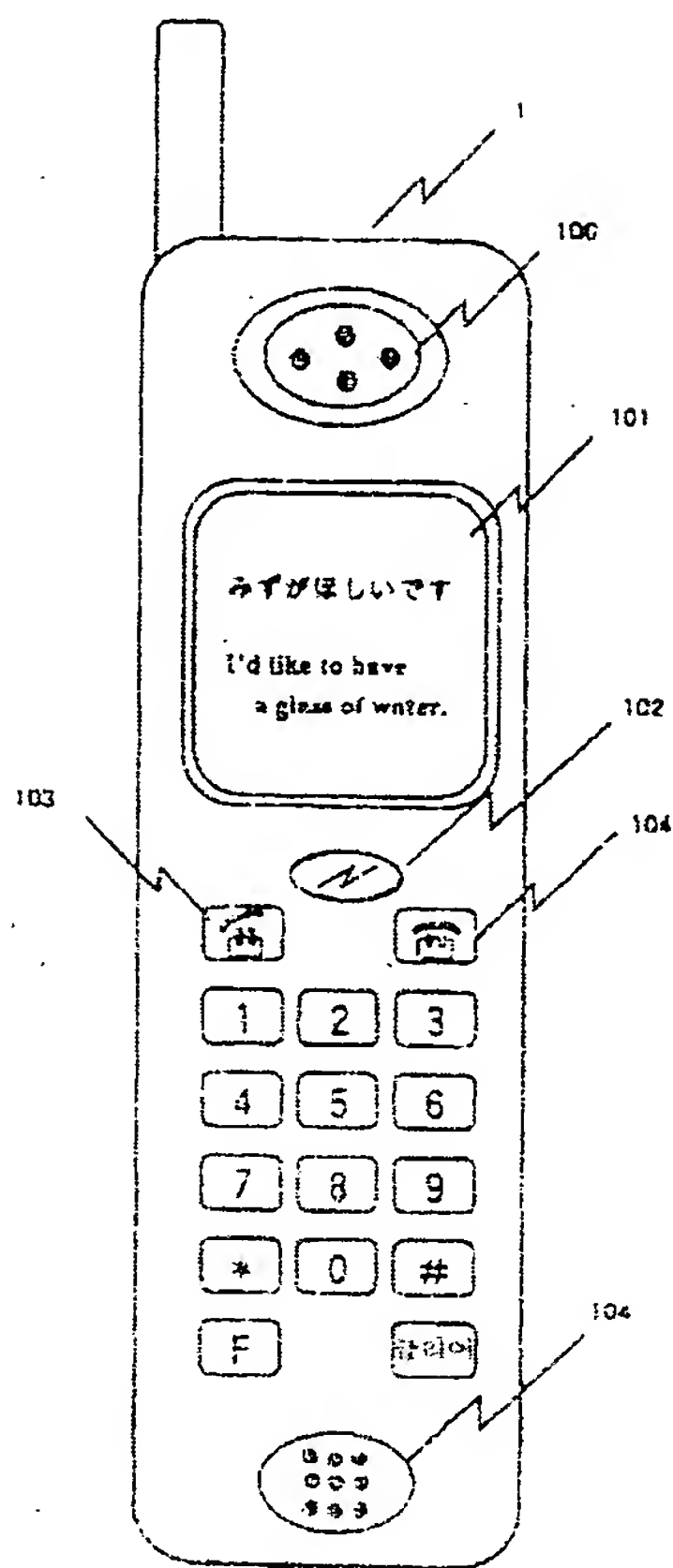
도면8



도면9

문제 번호	문장 번호	역어		
		1	2	...
:	:	:	:	:
4	1	...	I'd like to have a glass of water.	...
4	2	...	Can I see the menu?	...
4	3	...	I'd like to have a cup of coffee.	...
4	4	...	I'd like to have a spoon.	...
:	:	:	:	:

도면 10



도면 11

	사용자 ID	부과 요금
	:	:
401	1234	100
402	1235	0
403	1236	300
404	1237	350
	:	:



도면 12

	언어 종류 번호	언어 종류
801	1	일본어→한국어
802	2	일본어→영어
803	3	일본어→중국어
804	4	일본어→불어
	⋮	

도면 13.

	장면 번호	장면
901	1	공항
902	2	호텔
903	3	레스토랑
904	4	병원

도면 14

	장면 번호	문제 번호	문제
	1	1	[ ]はどこですか
	⋮	⋮	
501	3	1	こんにちは
502	3	2	ありがとう
503	3	3	[ ]はどこですか
504	3	4	[ ]が欲しいです
505	3	5	[ ]はありますか
506	3	6	[ ]てください
	⋮	⋮	

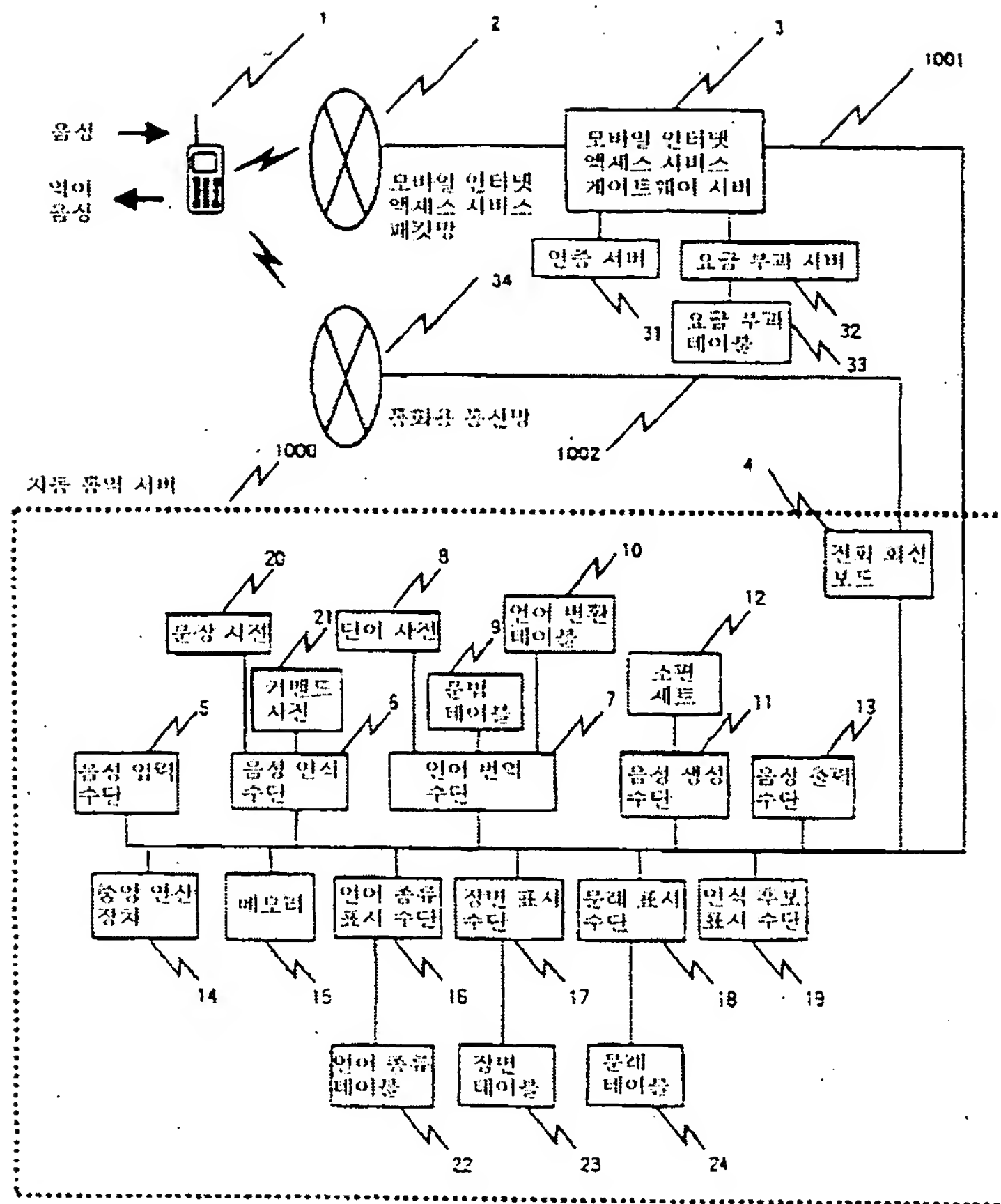
도면 15

	문례 번호	문장 번호	문장	특정 벡터 시계열
501	1	1	こんにちは	A11, A12, ...
602	2	1	ありがとう	A21, A22, ...
	:	:	:	:
603	4	1	みずがほしいです	A31, A32, ...
604	4	2	めにゃーがほしいです	A41, A42, ...
605	4	3	こーひーがほしいです	A51, A52, ...
606	4	4	すぶーんがほしいです	A61, A62, ...
	:	:	:	:

도면 16

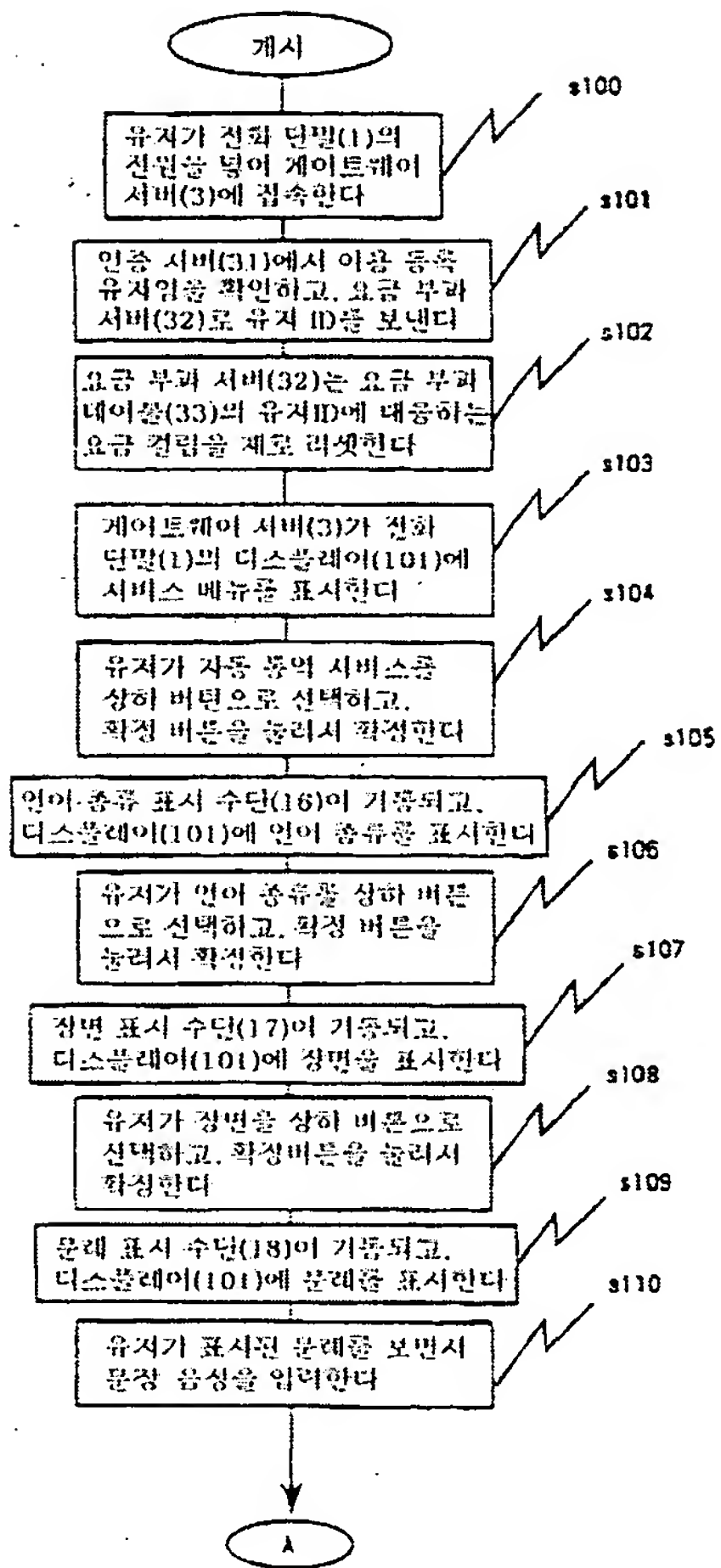
	커맨드 번호	커맨드	특정 벡터 시계열
701	1	かくてい	B11, B12, ...
702	2	つき	B21, B22, ...
703	3	サービス終了	:
	:	:	B31, B32, ...

도면 17

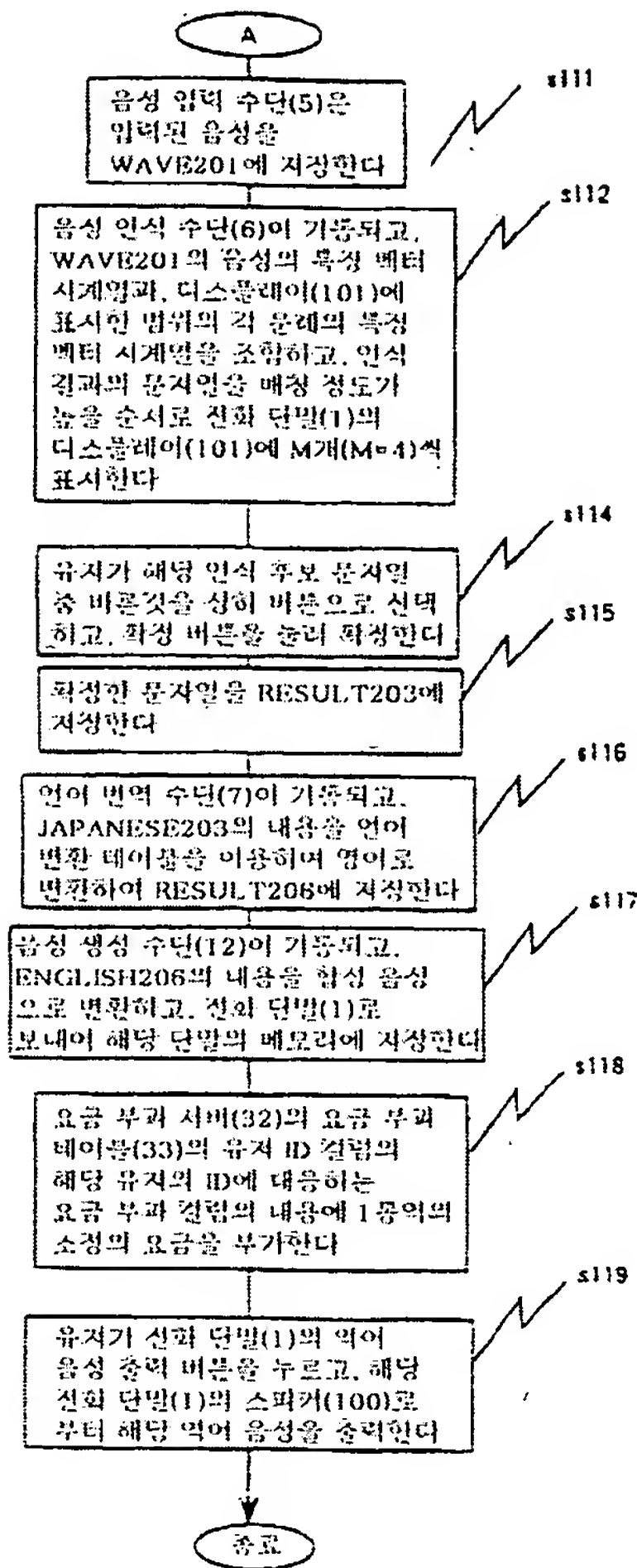




도면 18



도면 19



(71) 출원인가부시키가이샤 히타치세이사쿠쇼

일본 도쿄토 치요다쿠 간다스루가다이 4쵸메 6반치(72) 발명자

기따하라요시노리

일본도쿄도다찌까와시스나가와쵸7-20-36

오부찌야스나리

일본도쿄도고꾸분지시니시꼬이가꾸보3-8-1-62

고이즈미아쓰코

일본가나가와켄사가미하라시니시하시모토1-19-29

미즈타니세이끼

일본가나가와켄요코하마시아오바구스미레다이17-6-205(74) 대리인, 구영창, 장수길 **심사청구 : 있음(54)**  
**음성 통역 서비스 방법 및 음성 통역 서버요약**

본 발명의 목적은, 통화용 전화기, 특히, 보급율이 높은 모바일 인터넷 액세스 서비스 이용이 가능한 전화를 사용하여 통역 정확도가 높은 음성 통역을 누구나 간단히 이용할 수 있는 서비스를 제공하는 것에 있다. 상기 목적을 달성하기 위해, 본 발명에서는 유저가 전화를 이용하여 서버상에서 가동하는 자동 통역 서버에 음성을 입력한 후, 송신하면, 번역된 음성을 다시 유저의 전화에 회신하도록 한다. 그 때에, 최초에 유저가 모바일 인터넷 액세스 서비스 이용이 가능한 전화로부터 모바일 인터넷 액세스 서비스 패킷 망을 통해, 모바일 인터넷 액세스 서비스 게이트웨이 서버에 접속하면, 상기 자동 통역 서버는, 유저가 통역 가능한 언어 종류, 유저가 통역 가능한 문례(예시 문장 혹은 문장에) 장면, 유저가 입력 가능한 문례를, 상기 유저가 갖고 있는 전화의 표시 화면에 표시하도록 한다.

대표도

도 1

## 색인어

모바일 인터넷, 자동 음성 통역, 문례, 인증 서버, 통역 서버

## 명세서

### 도면의 간단한 설명

- 도 1은 본 발명의 제1 실시예를 도시하는 일영(日英) 자동 통역 서비스 시스템 구성도.
- 도 2는 메모리(15)의 데이터 구조의 예.
- 도 3은 전화 단말(1)의 예.
- 도 4는 전화 단말(1)의 디스플레이(101)에의 서비스 메뉴 표시예.
- 도 5는 전화 단말(1)의 디스플레이(101)에의 통역 가능한 언어 종류 표시예.
- 도 6은 전화 단말(1)의 디스플레이(101)에의 통역 가능한 장면 표시예.
- 도 7은 전화 단말(1)의 디스플레이(101)에의 통역 가능한 문례(文例) 표시예.
- 도 8은 전화 단말(1)의 디스플레이(101)에의 인식 결과 표시예.
- 도 9는 언어 변환 테이블(10)의 데이터 구조의 예.
- 도 10은 전화 단말(1)의 디스플레이(101)에의 통역 결과 표시예.
- 도 11은 요금 부과 테이블(33)의 데이터 구조의 예.
- 도 12는 언어 종류 테이블(22)의 데이터 구조의 예.
- 도 13은 장면 테이블(23)의 데이터 구조의 예.
- 도 14는 문례 테이블(24)의 데이터 구조의 예.
- 도 15는 문장 사전(20)의 데이터 구조의 예.
- 도 16은 커맨드 사전(21)의 데이터 구조의 예.
- 도 17은 본 발명의 제2 실시예를 도시하는 일영 자동 통역 서비스 시스템 구성도.
- 도 18은 제1 실시예의 동작을 도시하는 흐름도.
- 도 19는 제1 실시예의 동작을 도시하는 흐름도.

(도면의 주요 부분에 대한 부호의 설명)

- 1 : 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말
- 2 : 모바일 인터넷 액세스 서비스 패킷 망
- 3 : 모바일 인터넷 액세스 서비스 게이트웨이 서버
- 4 : 전화 회선 보드
- 5 : 음성 입력 수단
- 6 : 음성 인식 수단
- 7 : 언어 번역 수단
- 8 : 단어 사전
- 9 : 문법 테이블
- 10 : 언어 변환 테이블
- 11 : 음성 생성 수단
- 12 : 소편 세트(음성 분절 세트)
- 13 : 음성 출력 수단
- 14 : 증강 연산 장치
- 15 : 메모리
- 16 : 언어 종류 표시 수단
- 17 : 장면 표시 수단

- 18 : 문례(예시 문장) 표시 수단
- 19 : 인식 후보 표시 수단
- 31 : 인증 서버
- 32 : 요금 부과 서버
- 33 : 요금 부과 테이블

## 발명의 상세한 설명

### 발명의 목적

#### 발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 사용자가 발성한 음성을 다른 언어로 번역하여 음성으로서 출력하는 자동 통역 서비스에 관한 것이다. 주로 모바일 인터넷 액세스 서비스 이용이 가능한 휴대 전화를 이용한 서비스를 대상으로 하지만, 통상의 유선 전화를 이용하거나, 퍼스널 컴퓨터, 휴대 정보 단말 등을 이용하여 상기 동작을 행하는 음성 통역 서비스도 포함한다.

종래, 음성 통역기라는 것이 있었다. 예를 들면, 특개평 11-125539에는, 내장 마이크로폰으로부터 입력되어 음성을 다른 언어로 번역하고, 번역 결과를 마이크로폰과는 반대 면에 설치된 스피커로부터 음성 출력하는, 소형이면서 또한 한 손 조작이 가능한 음성 통역 장치가 개시되어 있다. 그러나, 이러한 음성 통역 장치는 전용기(專用機)이기 때문에, 다른 용도로는 이용할 수 없고, 해외 여행 등의 실제 이용 시에는, 결국 유저의 소지품을 늘려 버리는 것이 된다. 한편, 예를 들면, 특개평9-65424에는, 음성 인식 서버와 무선 휴대 단말을 이용한 음성 번역 시스템이 개시되어 있다. 그러나, 이 음성 번역 시스템에서는, 임의 문장을 입력 대상으로 하고 있기 때문에, 높은 통역 정밀도가 얻어지기 어렵고, 실제 운용이 곤란하다는 문제가 있었다.

#### 발명이 이루고자하는 기술적 과제

본 발명은 상기한 점을 감안하여 이루어진 것으로, 통화용 전화기, 특히, 보급율이 높은 모바일 인터넷 액세스 서비스 이용이 가능한 전화를 이용하여, 정밀도가 높은 음성 통역을 누구나 간단히 이용할 수 있는 서비스를 제공하는 데에 그 목적이 있다.

### 발명의 구성 및 작용

본 발명에서는, 유저가 전화를 이용하여 서버상에서 가동되는 자동 통역 서버에 음성을 송신하면, 번역된 음성이, 다시 유저의 전화에 회신되도록 한다. 그때에, 최초에 유저가 모바일 인터넷 액세스 서비스 이용이 가능한 전화로부터 모바일 인터넷 액세스 서비스 패킷망을 통하여, 모바일 인터넷 액세스 서비스 게이트웨이 서버에 접속하면, 상기 자동 통역 서버가, 유저가 통역 가능한 언어 종류를, 유저가 갖고 있는 전화의 표시 화면에 메뉴로서 표시하여, 유저로 하여금 상기 언어 종류 메뉴로부터 통역 희망 언어를 선택할 수 있도록 한다. 또한, 유저가 통역 가능한 문례 장면을 상기 유저의 전화의 표시 화면에 표시하여, 유저가 상기 장면 메뉴로부터 통역 가능 문례 장면을 선택할 수 있도록 한다. 또한, 유저가 입력 가능한 문례를 상기 유저의 전화의 표시 화면에 표시하여, 유저가 상기 입력 가능한 문례를 보면서 발성 입력을 할 수 있도록 한다. 또한, 상기 자동 통역 서버는, 상기 입력된 음성을 상기 표시한 범위의 문례 사진을 이용하여 인식하고, 역어(譯語)로 변환하여, 음성으로서 상기 전화 단말에 출력한다.

이와 같이, 새롭게 전용기를 구입하지 않고 일반 전화 단말을 이용하여, 상기 전화 단말에 표시한 정형적(定型的)인 문장을 입력 대상으로 하여, 자동 통역 서버로 통역 처리를 함으로써, 통역 정확도가 높은 통역 서비스를 용이하게 받을 수 있는 환경을 제공할 수가 있다.

### (실시의 형태)

이하, 본 발명의 실시예를 도면을 이용하여 설명한다.

도 1은 본 발명의 제1 실시예를 도시하는 자동 통역 서비스 시스템의 구성도이다. 본 발명은, 음성 통역기에 관한 것으로, 본 실시예에 있어서의 서버는, 퍼스널 컴퓨터, 워크스테이션 등, 중앙 연산 장치, 메모리를 구비하며, 동도와 같은 구성을 할 수 있는 것이거나 좋고, 플랫폼의 종류가 본 발명의 적용 범위를 한정하는 것은 아니다. 또한, 본 실시예는 영어를 일본어로 통역하는 자동 통역 서비스이지만, 이것은 일례이고, 일영, 중일, 일중 등, 언어의 종류는 제한되지 않는다.

동 도면에 있어서, 참조 부호 1은 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말, 2는 모바일 인터넷 액세스 서비스 패킷망, 3은 모바일 인터넷 액세스 서비스 게이트웨이 서버, 4는 전화 회선 보드, 5는 음성 입력 수단, 6은 음성 인식 수단, 7은 언어 번역 수단, 8은 단어 사전, 9는 문법 테이블, 10은 언어 변환 테이블, 11은 음성 생성 수단, 12는 소편(素片) 세트(speech segments set: 음성 문절 세트), 13은 음성 출력 수단, 14는 중앙 연산 장치, 15는 메모리, 16은 언어 종류 표시 수단, 17은 장면 표시 수단, 18은 문례 표시 수단, 19는 인식 후보 표시 수단, 20은 문장 사전, 21은 커맨드 사전, 22는 언어 종류 테이블, 23은 장면 테이블, 24는 문례 테이블, 31은 인증 서버, 32는 요금 부과 서버, 33은 요금 부과 테이블이다. 메모리(15)의 데이터 구조를 도 2에 도시한다. 또한, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)의 개관의 예를 도 3에 도시한다. 또, 여기서는, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말로서, 무선, 통화 음성 및 데이터를 동일한 프로토콜로 다룰 수 있는

전화 단말을 가정하여 설명한다.

도 1, 도 3에 있어서, 우선, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)의 전원(104)을 넣어, 모바일 인터넷 접속용 버튼(102)을 누르면, 모바일 인터넷 액세스 서비스 패킷망(2)을 통해, 센터의 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에 접속되고, 인증 서버(31)에서 모바일 인터넷 액세스 서비스 이용 등록 유저인 것이 확인된다. 이 시점에서, 유저 ID가 요금 부과 서버(32)에 보내어진다. 당연하지만, 유저 ID는 전화 단말(1)의 ID와 링크하고 있는 것이 많지만, 전화 단말(1)과는 무관하게 유저 ID를 패스워드와 조합하여 입력시키는 방법도 있다.

요금 부과 서버(32)는 요금 부과 테이블(33)을 갖는다. 요금 부과 테이블(33)의 데이터 구조를 도 11에 도시한다. 인증 서버(31)로부터 보내어진 ID는, 요금 부과 테이블(33)의 유저 ID 컬럼(411)의 각 셀과 대조되어, 일치한 셀의 요금 컬럼(412)이 제로 리세트된다. 예를 들면, 상기 유저의 유저 ID가 '1236'이라고 하면, 요금 부과 테이블(33)의 유저 ID 컬럼(411)의 셀 중 '1236'인 셀(403)과 일치하기 때문에, 요금 컬럼(412)의 403에 대응하는 셀이 '0'이 된다.

모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에는, 본 실시예의 자동 통역 서버와 같은 각종 서비스 서버가 전용선으로 접속되어 있거나, 일반 인터넷 망이 접속되어 있다. 또, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에의 접속에는, 상기와 같은 모바일 인터넷 접속용 버튼(102)을 누르는 방법 외에, 어떤 소정의 전화번호를 숫자 버튼으로 입력하는 방법 등을 이용하여도 좋다.

모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)이, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에 접속되어, 인증 서버(31)에서 모바일 인터넷 액세스 서비스 이용 등록 유저인 것이 확인되면, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)는, 전화 단말(1)의 디스플레이(101)에, 도 4와 같이 서비스의 메뉴를 표시한다. 상기 서비스 메뉴는, 도 4와 같이, 흑백 반전에 의해 1번째 항목의 휘도가 다른 항목과는 다른 상태에서 초기 표시된다. 이것은, 제1번째의 항목이 선택되어 있는 것을 나타내지만, 물론, 그 표시는 흑백 반전에 한하지 않고, 다른 항목과 구별이 되어질 수 있는 방법이면 좋다.

유저는, 상기 서비스 메뉴를 보면서 전화 단말(1)상의 상하 커서 이동의 기능이 할당된 소정의 버튼을 눌러, 흑백 반전을 제3번째의 항목 '자동 통역'까지 이동시키고, 또한 전화 단말(1)상의 상기 소정의 버튼과는 다른 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 메뉴 항목 '자동 통역'이 확정되면, 전화 단말(1)이 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)를 경유하여 자동 통역 서버(1000)에 접속된다.

다음에, 자동 통역 서버(1000)의 언어 종류 표시 수단(16)이 기동되고, 언어 종류 테이블(22)을 이용하여 전화 단말(1)상의 디스플레이(101)에, 도 5와 같이 통역 가능한 언어의 종류를 표시한다. 언어 종류 테이블(22)은 도 12와 같은 데이터 구조를 갖고 있고, 언어 종류 표시 수단(16)은 상기 테이블의 언어 종류(12)의 각 항목을, 상기 전화 단말(1)에 보내어 도 5와 같이 상기 전화 단말(1)상의 디스플레이(101)에 표시한다. 동 도면에서는, 흑백 반전에 의해 1번째의 항목의 휘도가 다른 항목과는 다른 상태에서 초기 표시되어 있다. 이것은, 제1번째의 항목이 선택되어 있는 것을 나타내지만, 물론, 그 표시는 흑백 반전에 한하지 않고, 다른 항목과 구별이 되어질 수 있는 것 같은 방법이면 좋다. 유저는, 상기 언어 종류 메뉴를 보면서 전화 단말(1)상의 상하 커서 이동의 기능을 할당한 소정의 버튼을 눌러, 흑백 반전을 예를 들면 제2번째의 항목 '일본어 → 영어'까지 이동시키고, 또한 전화 단말(1)상의 상기 소정의 버튼과는 다른 소정의 확정 기능이 할당된 버튼을 눌러 확정한다. 이때, 언어 종류 표시 수단(16)은 전화 단말(1)의 커서의 위치를 수취하여 메모리(15)상의 LANG209에 상기 위치 번호를 저장한다. 예를 들면, 일본어를 영어로 통역하고 싶은 경우에는, 「일본어 → 영어」가 2번째에 있으므로, 메모리(15)상의 LANG209에 '2'가 저장된다.

물론, 이 언어 종류 지정은, 전화 단말(1)상의 디스플레이(101)에 언어 종류를 표시하여, 유저에게 원하는 언어 종류를 상하 이동 버튼으로 선택시키는 방법 외에, 상기 전화 단말(1)상의 디스플레이(101)를 터치 패널 디스플레이로 구성하고 손가락 등의 접촉에 의해 원하는 언어 종류를 지정하는 방법을 이용하도록 해도 좋다.

또한, 미리 어느 소정의 전화번호를 언어 종류마다 할당해 놓고, 유저가 원하는 언어 종류에 대응하는 전화번호를 상기 전화 단말(1)상의 숫자 버튼으로 입력하는 방법 등을 이용하여도 좋다.

언어 종류가 확정되면, 다음에, 자동 통역 서버(1000)의 장면 표시 수단(17)이 기동되고, 장면 테이블(23)을 이용하여, 전화 단말(1)상의 디스플레이(101)에, 도 6과 같이 통역 가능한 장면을 표시한다. 여기서 말하는 장면이란, 「공항」, 「호텔」, 「레스토랑」 등, 본 발명의 통역 서비스를 이용하는 장면의 것이다. 장면 테이블(23)은, 도 13과 같은 데이터 구조를 갖고 있고, 장면 표시 수단(17)은 상기 테이블의 장면(912)의 각 항목을, 상기 전화 단말(1)에 보내어 도 6과 같이 상기 전화 단말(1)상의 디스플레이(101)에 표시한다. 도 13에서는, 흑백 반전에 의해 1번째의 항목의 휘도가 다른 항목과는 다른 상태에서 초기 표시되어 있다. 이것은, 제1번째의 항목이 선택되어 있는 것을 나타내지만, 물론, 그 표시는 흑백 반전에 한하지 않고, 다른 항목과 구별짓는 방법이면 좋다.

유저는, 상기 장면 메뉴를 보면서 전화 단말(1)상의 상하 커서 이동의 기능이 할당된 소정의 버튼을 눌러, 흑백 반전을 예를 들면 제3번째의 항목 '레스토랑'까지 이동시키고, 또한 전화 단말(1)상의 상기 소정의 버튼과는 다른 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 이때, 장면 표시 수단(17)은 전화 단말(1)의 커서의 위치를 수취하여 메모리(15)상의 SCENE210에 상태당 위치 번호를 저장한다. 예를 들면, 레스토랑의 장면으로 통역하고 싶은 경우에는, 「레스토랑」이 3번째에 있으므로, 메모리(15)상의 SCENE210에 '3'이 저장된다.

물론, 이 장면 지정은, 전화 단말(1)상의 디스플레이(101)에 장면을 표시하여, 유저에게 원하는 장면을 상하 이동 버튼으로 선택시키는 방법 외에, 상기 전화 단말(1)상의 디스플레이(101)를 터치 패널 디스플레이로 구성하고 손가락 등의 접촉에 의해 원하는 장면을 지정하는 방법을 이용하도록 해도 좋다. 또한, 미리 있는 소정의 전화번호를 장면마다 할당해 놓고, 유저가 원하는 장면에 대응하는 전화번호를 상기 전화 단말(1)상의 숫자 버튼으로 입력하는 방법 등을 이용하여도 좋다.



장면이 확정되면, 다음에, 자동 통역 서버(1000)의 문례 표시 수단(18)이 기동되고, 문례 테이블(24)을 이용하여 전화 단말(1)상의 디스플레이(101)에, 도 7과 같이 통역 가능한 문례를 표시함과 동시에, 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된다. 계속해서, 음성 입력 수단(5)은, 시스템을 음성 입력 가능한 상태로 한다. 그후, 유저는, 상기 문례를 보면서, 예를 들면 전번의 예이면, 일본어로 레스토랑의 장면에서의 통역시키고 싶은 문장을 전화 단말(1)의 송화기의 마이크(104)를 향해서 발성 입력한다. 여기서, 예를, 레스토랑의 장면에서 「물을 원합니다。」라는 문장을 영어로 통역하고 싶은 경우를 생각하는 것으로 한다.

문례 테이블(24)은 도 14와 같은 데이터 구조를 갖고 있고, 문례 표시 수단(18)은, 문례 테이블(24)의 장면 번호(511)의 각 항목 중, 메모리(15)상의 SECENE210에 저장되어 있는 값의 항목의 문례(513)를 문례 번호(512)의 '1'로부터 순차적으로, 상기 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 표시한다. 전번의 예에서는, 메모리(15)상의 SCENE210에 '3'이 저장되어 있기 때문에, 도 14의 문례 테이블(24)의 장면 번호(511)의 값이 '3'인 항목(501, 502, 503, 504)의 문례(513)의 내용 「안녕하십니까」, 「고맙습니다」, 「[ ]는 어디입니까」, 「[ ]를 원합니다」를 상기 전화 단말(1)에 보내어 도 7과 같이 상기 전화 단말(1)상의 디스플레이(101)에 M개씩 순차 표시한다. M은 디스플레이(101)의 크기에 따라서 설정하는 상수이지만, 여기서는 M = 4으로 한다.

유저는, 상기 문례의 속에 「[ ]를 원합니다」라는 문례가 있기 때문에, 이 문례를 기초로 「물을 원합니다」로 발성 입력한다. 이 발성 입력에 있어서, 전화 단말(1)상의 발성 입력의 신호의 의미를 가지게 한 소정의 버튼을 눌러 자동 통역 서버(1000)의 음성 입력 수단(5)을 음성 입력 가능한 상태로 하고 나서 발성해도 좋고, 특히 버튼 등의 트리거를 이용하지 않고 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된 후 언제나 음성 입력 가능한 상태로 해도 좋다. 또, 표시하는 문례는, 상술한 바와 같이 슬롯 [ ]을 함유한 문장 혹은 문법이라도 좋고, 혹은 문장 그 자체라도 좋다. 여기서, 슬롯이란, 단어, 구 등을 그 장소에 설정할 수 있다고 하는 말하자면 박스이다. 예를 들면, 「[ ]를 원합니다」에서는, 슬롯 [ ]에 「물」, 「커피」, 「찬 물」 등을 설정할 수 있다.

이와 같이, 문례를 표시하여 문형을 어느 정도 규정함으로써, 음성 인식의 정확도를 현저히 향상시킬 수 있다. 또한, 문례를 표시함으로써 유저의 사용성을 향상시킬 수 있다.

또, 상기 표시 문례는, 상하 커서 이동의 기능을 할당한 소정의 버튼을 누름으로써, 차례 차례로 스크롤하고 인식 가능한 문례를 전부 순차 표시시킬 수 있다. 다만, 이때에, 상기 전화 단말(1)상의 디스플레이(101)에 표시시키고 있는 최초의 문례의 문례 번호(512)의 값, 및, 표시시키고 있는 최후의 문례의 문례 번호(512)의 값을 메모리(15)상의 BSENTENCE211, ESENTENCE212에 각각 저장한다. 도 7의 예에서는, BSENTENCE211에는 '1', ESENTENCE212에는 '4'가 저장된다.

계속해서, 음성 입력 수단(5)은, 전화 회선 보드(4)로서 아날로그/디지털 변환된 입력 음성을 메모리(15)상의 WAVE201에 저장한다. 전화 회선 보드(4)에서의 아날로그/디지털 변환의 샘플링률은 8 kHz, 11 kHz, 16 kHz 등, 유저가 적절하게 정할 수 있다.

또, 이 시점에서, 상기 유저가, 입력한 음성을 취소하고, 재 입력하고 싶다고 하면, 전화 단말(1)상의 취소 기능이 할당된 소정의 버튼을 누름으로써, 음성을 재 입력할 수 있는 상태로 할 수 있게 된다. 전화 단말(1)상의 취소 기능이 할당된 소정의 버튼은, 누르면, 발성 입력의 신호의 의미를 가진 소정의 버튼을 눌렀을 때 혹은 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된 직후와 동일한 초기 상태로 리셋되어, 음성 입력 가능한 상태가 되도록 한다.

다음에, 음성 인식 수단(6)이 기동된다. 음성 인식 수단(6)은, 메모리(15)상 WAVE201에 저장된 음성 데이터를 읽어들이, 특징 벡터의 시계열로 변환하고, 각 문장(sentence) 음성의 상기 특징 벡터 시계열을 갖는 문장 사전(20)을 이용하여 대조를 행하고, 상기 음성 데이터를 인식하여 후보를 출력한다. 특징 벡터 시계열에의 변환 방법이나, 대조 방법을 포함해서 음성을 인식하는 방법으로서, 예를 들면 L. Rabiner & B. -H. Juang저, 古井貞熙 監譯 「음성 인식의 기초(하)」 (NTT 어드밴스 테크놀로지, 1995) PAGE 245~P304 기재의 방법 등을 이용할 수가 있다. 물론, 다른 음성 인식방법을 이용하여도 좋고, 음성 인식 방법이 본 발명을 한정하는 것은 아니다. 문장 사전(20)의 데이터 구조를 도 15에 도시한다. 음성 인식 수단(6)은, 메모리(15)상 WAVE201에 저장된 음성 데이터를 읽어들이, 문장 사전(20)의 문례 번호(611)의 값이 메모리(15)상의 BSENTENCE211, ESENTENCE212에 각각 저장되어 있는 값의 범위내인 항목의 것 전체의 특징 벡터 시계열(614)의 값을 이용하여 음성 인식을 행한다. 전번의 예에서는, BSENTENCE211에는 '1', ESENTENCE212에는 '4'가 저장되어 있기때문에, 문장 사전(20)의 문례 번호611의 값이 '1'에서 '4'까지의 각 항목의 특징 벡터 시계열(614)의 값을 이용하여 음성 인식을 행한다. 그 결과, 상기 음성은, 매칭 정보가 높은 순서로, 예를 들면, 문자열 '메뉴를 원합니다', '물을 원합니다', '커피를 원합니다', '스폰을 원합니다'의 문례 번호 및 문장 번호로 변환된다. 그 때문에, 이들 후보의 문례 번호611, 문장 번호612, 및 문장613을 매칭도가 높은 순서로, 메모리(15)상의 RECOGNUM(1), RECOGSNUM(1), RECOGS(1), RECOGNUM(2), RECOGS(2), ..., RECOGNUM(N), RECOGSNUM(N), RECOGS(N)205에 저장하여 간다. 여기서, N은, 문장 사전(20)의 문례 번호611의 값이 메모리(15)상의 BSENTENCE211, ESENTENCE212에 각각 저장되어 있는 값의 범위내인 항목의 전체 수이다. 다음에, 인식 후보 표시 수단(19)이 기동되어, RECOGS(1), RECOGS(2), ..., RECOGS(M) 205의 내용을, 도 8과 같이, 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순서대로 표시한다. 이때, 메모리(15)상의 ICNT204에 '1'을 저장하고, 전화 단말(1) 상의 디스플레이(101)에 표시되어 있는 RECOGS(1CNT)의 내용이 흑백 반전되도록 한다. M은 디스플레이(101)의 크기에 따라서 설정하는 상수이지만, 여기서는 M = 4로 한다. 또한, 메모리(15)상의 INDEX215에 '0'을 저장한다.

유저는, 제1 후보의 표시가 자신이 발성한 내용과 일치하고 있는지 혹은 근접하다면, 그대로 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 제1 후보 표시가 정확하지 않은 경우, 유저는, 상하 커서 이동의 기능이 할당된 소정의 버튼을 누름으로써, 전화 단말(1)상의 디스플레이(101)상의 정확한 문자열이 표시되어 있는 부분에까지 흑백 반전 표시 영역을 아래쪽으로 이동시킨다. 즉, 하 이동 버튼을 누를 때마다, 메모리(15)상의 ICNT 204의 값을 인크리먼트하고, 전화 단말(1)상의 디스플레이(101)상에 동일하게 메모리(15)상의 RECOG(ICNT)의 내용이 표시되어 있는 부분만을 흑백 반전 표시하도록 한다. 여기

서, ICNT 204의 값이 M을 넘은 경우에는, 메모리(15)상의 INDEX215의 값에 'M'을 가산하고, 다음 M 개의 후보 RECOGS(INDEX+1), RECOGS(INDEX+2), ..., RECOGS(INDEX+M)를 메모리(15)상의 가운데부터 판독하며, 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차 표시한다. 이때, 메모리(15)상의 ICNT204에 '1'을 저장하고, 전화 단말(1)상의 디스플레이(101)에 표시되어 있는 RECOGS(INDEX+1), RECOGS(INDEX+2), ..., RECOGS(INDEX+M)중의 ICNT 번째의 표시가 흑백 반전되도록 한다. 이하, 마찬가지로 하여, 다음의 M 개의 후보를, 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차적으로 표시해 간다. 또, 도 15에 도시한 음성 인식에 사용하는 문장 사전(20)의 구조는 일례이고, 예를 들면, 문법과 단어 사전을 조합하는 방법도 있다. 그 경우에도, 메모리(15)상의 SCENE210에 저장된 장면 번호에 대응시켜, 장면별 문법이나 사전을 이용하도록 한다.

또한, 상 이동 버튼으로 흑백 반전 표시 영역을 위쪽으로 이동시키도록 하고, 백하여 표시 내용을 보이도록 해도 좋다. 즉, 상 이동 버튼을 누를 때마다, 메모리(15)상의 ICNT 204의 값을 디크리먼트하여, 전화 단말(1)상의 디스플레이(101)상의 RECOGS(INDEX+1), RECOGS(INDEX+2), ..., RECOGS(INDEX+M)중의 ICNT 번째의 표시부분만을 흑백 반전 표시하도록 한다.

이 후보의 선택은, 상기 전화 단말(1)상의 상하 이동 버튼을 이용하는 방법에 한하지 않고, 상기 전화 단말(1)상의 디스플레이(101)를 터치 패널 디스플레이로 구성하고 손가락 등의 접촉에 의해 후보를 지정하는 방법을 이용하도록 해도 좋다.

유저가, 흑백 반전되어 있는 표시가 자신이 발성한 내용과 일치하고 있거나 혹은 가까운 것으로 하여 확정 기능이 할당된 소정의 버튼을 눌러 확정하면, 메모리(15)상의 RECOGNUM(INDEX+ICNT), RECOGSNUM(INDEX+ICNT), RECOGS(INDEX+ICNT)의 값을, 동일하게 메모리(15)상의 PCAND213, SCAND214, JAPANESE203에 각각 저장한다.

도 8의 예에서는, '물을 원합니다'가, 제2행째에 표시되어 있기 때문에, 하 이동 버튼을 누름으로써 제2행째까지, 흑백 반전 영역을 이동시키고, 확정 버튼을 누른다. 그리고, INDEX는 '0' ICNT는 '2'가 되고, RECOGNUM(2), RECOGSNUM(2), RECOGS(2)의 값인 '4', '1', '물을 원합니다'가 메모리(15)상의 PCAND213, SCAND214, JAPANESE203에 각각 저장된다.

자신이 발성한 내용의 확인 방법에는, 이상 진술한 바와 같은 전화 단말(1)상의 디스플레이(101)상에 음성 인식 후보를 표시하는 방법 외에, 디스플레이(101)를 사용하지 않고서, 이하와 같은 방법을 이용할 수도 있다. 즉, 음성 인식 수단(6)이 매칭 정도가 높은 순으로 후보의 문례 번호(611), 문장 번호(612), 및 문(613)을 매칭 정도가 높은 순으로, 메모리(15)상의 RECOGNUM(1), RECOGSNUM(1), RECOGS(1), RECOGNUM(2), RECOGSNUM(2), RECOGS(2), ..., RECOGNUM(N), RECOGSNUM(N), RECOGS(N)205에 저장한 후, 음성 생성 수단(12)이 기동되어, 이때, 메모리(15)상의 JCNT208에 '1'을 저장하고, 메모리(15)상의 RECOGS(JCNT)를 먼저 판독하여, 상기 문자열을 합성 음성으로 변환하고, 상기 음성의 파형 데이터를 디지털/아날로그 변환에 의해 아날로그 데이터로 변환하여, 음성 출력 수단(13)을 통하여 음성으로서 전화 단말(1)에 보낸다. 문자열을 합성 음성으로 변환하기 위해서는, 예를 들면, J.Allen, M.S.Hunnicuttt and D.Klatt 등의 'From Text to Speech'(Cambridge University Press, 1987) P.16에서 P.150에 기재되어 있는 홀만트 합성 방식이나, 失頭の 「텍스트 음성 합성 기술의 최신 상황」(Interface, Dec., 1996) P.161내지 P.165에 기재되어 있는 파형 중첩 방식을 쓸 수 있다. 물론, 다른 텍스트 음성 합성 방식을 이용할 수도 있다. 또한, 인식 가능한 문례에 대응하는 음성을 미리 녹음하여, 메모리 등의 기억 장치에 저장하여 놓도록 해도 좋다.

유저는, 상기 전송된 음성을 전화 단말(1)상의 스피커(100)로부터 듣고, 상기 음성이 자신이 발성 입력한 내용과 합치하고 있으면, 전화 단말 상의 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 상기 음성이 자신이 발성 입력한 내용과 합치하지 않는 경우에는, 유저는 전화 단말(1)상의 다음 후보 제시의 기능이 할당된 소정의 버튼을 누르도록 한다. 자동 통역 서버(1000)중의 음성 생성 수단(12)은, 전화 단말(1)상의 다음 후보 제시 기능이 할당된 소정의 버튼이 눌리면, 메모리(15)상의 JCNT208을 인크리먼트하여, RECOGS(JCNT)를 판독하고, 상기 문자열을 합성 음성으로 변환하며, 상기 음성의 파형 데이터를 디지털/아날로그 변환에 의해 아날로그 데이터로 변환하여, 음성 출력 수단(13)을 통해 음성으로서 전화 단말(1)에 보낸다.

상술한 동작과 마찬가지로, 유저는, 상기 전송된 음성을 전화 단말(1)상의 스피커(100)로부터 듣고, 상기 음성이 자신이 발성 입력한 내용과 합치하고 있으면, 전화 단말(1)상의 확정 기능이 할당된 소정의 버튼을 눌러 확정하고, 상기 음성이 자신이 발성 입력한 내용과 합치하지 않으면, 유저는 전화 단말(1)상의 다음 후보 제시의 기능이 할당된 소정의 버튼을 눌러, 자신이 발성 입력한 내용에 합치하는 음성이 들릴 때까지, 이상의 동작을 반복한다.

확정 버튼을 누르면, 메모리(15)상의 RECOGS(ICNT)에 저장되어 있는 문자열을 동일하게 메모리(15)상의 JAPANESE 203에 저장한다. 이 확정 버튼을 누르는 대신에, 어느 소정의 단어 혹은 구(句) 혹은 문장을 유저가 발성 입력하는 방법을 이용해도 좋다. 즉, 유저는, 상기 전송된 음성을 전화 단말(1)상의 스피커(100)로부터 듣고, 상기 음성이 자신이 발성 입력한 내용과 합치하고 있으면, 합치하고 있다는 취지의 소정의 단어 혹은 구 혹은 문장을 확정 음성으로서 전화 단말(1)상의 마이크(104)로부터 발성 입력하고, 상기 음성이 자신이 발성 입력한 내용과 합치하지 않으면, 합치하고 있는 경우와는 다른 소정의 단어 혹은 구 혹은 문장을 비확정 음성으로서 전화 단말(1)상의 마이크(104)로부터 발성 입력하도록 하여, 자동 통역 서버(1000)중의 음성 인식 수단(6)이, 상술한 바와 같은 문장 입력과 마찬가지로 방법으로 음성을 인식하여, 전체 매칭 정도가 미리 설정한 임계치를 하회하는 경우 혹은 ICNT204의 값이 N을 넘은 경우, 커맨드 사전(21)과의 대조를 개시한다. 커맨드 사전(21)의 데이터 구조를 도 16에 도시한다. 입력 음성의 특징 벡터의 시계열을, 커맨드 사전(21)의 각 항목의 특징 벡터의 시계열과 대조하여, 매칭 정도가 가장 높은 후보의 커맨드 번호를 갖고 커맨드로 한다. 여기서는 예로서, 유저가 「확정」이라고 발성 입력한 경우를 상정하면, 먼저, 문장 사전(20)을 이용한 인식으로, 상기 음성의 특징 벡터와 각 항목의 특징 벡터의 대조로 매칭 정도가 미리 설정한 임계치를 하회하는 것이 되고, 커맨드 사전(21)의 각 항목의 특징 벡터의 대조가 행하여져서, 701의 항목이 인식 후보가 된다. 상기 항목의 커맨드 번호가



10이면, 그것은 「확정」을 나타내는 발성 입력이 된다.

확정 음성이면, 메모리(15)상의 RECOGS(ICNT)에 저장되어 있는 문자열을 동일하게 메모리(15)상의 JAPANESE203에 저장한다. 또한, 비확정 음성이면, 메모리(15)상의 JCNT208을 인크리먼트하고, RECOGS(JCNT)를 판독하여, 상기 문자열을 합성 음성으로 변환하며, 상기 음성의 파형 데이터를 디지털/아날로그 변환에 의해 아날로그 데이터로 변환하여, 음성 출력 수단(13)을 통해 음성으로서 전화 단말(1)에 보낸다고 하는 상술한 동작을 확정 음성이 될 때까지 반복하도록 한다.

다음에, 자동 통역 서버(1000) 중의 언어 번역 수단(7)이 기동된다. 언어 번역 수단(7)은, 언어 변환 테이블(10)을 이용하여 메모리상의 JAPANESE203에 저장되어 있는 문자열을 다른 언어로 번역 변환한다. 이하, 언어 번역 수단(7)의 동작에 관해서 설명한다. 언어 변환 테이블(10)의 데이터 구조를 도 9에 도시한다.

우선, 언어 번역 수단(7)은, 메모리(15)상의 PCAND213, SCAND214에 저장되어 있는 값을 언어 변환 테이블(10)의 문례 번호311 및 문장 번호312의 각 항목과 순차 대조하여, 일치한 항목의 역어312 중의 LANG209의 값의 컬럼의 내용을 메모리(15)의 RESULT206에 저장한다.

앞의 예에서는, PCAND213, SCAND214에 저장되어 있는 값은 각각 '4', '1'이므로, 303의 항목 '물을 원합니다'와 일치한다. 또한, LANG209의 값은 '2'이기 때문에, 대응하는 역어312의 'I'd like to have a glass of water'를 메모리(15)의 RESULT206에 저장한다. 역어로의 변환에 대해서는, 상술한 예와 같은 언어 변환 테이블을 이용하는 방법에만 한하지 않고, 예를 들면 특개평03-328585에 기재되어 있는 용례 데이터 베이스를 이용한 번역 방법이나, 특개평03-51022에 기재되어 있는 구문 해석에 의한 번역 방법 등을 사용하더라도 좋다.

다음에, 언어 번역 수단(7)은, 도 10에 도시한 바와 같이, 메모리(15)의 JAPANESE 203 및 RESULT 206에 각각 저장되어 있는 내용을 전화 단말(1)상의 디스플레이(101)에 표시한다. 등 도면의 표시 방법은 일례이고, 표시 방법이 본 발명의 범위를 한정하는 것은 아니다.

계속해서, 자동 통역 서버(1000) 중의 음성 생성 수단(12)이 기동된다. 음성 생성 수단(12)은 메모리(15)상의 ENGLISH 206에 저장되어 있는 문자열을 읽어들이고, 상기 문자열을 합성 음성으로 변환하고, 메모리(15)상의 SYNWAVE 207에 파형 데이터를 저장한다. 문자열을 합성 음성으로 변환하기 위해서는, '예를 들면, J.Allen, M.S.Hunnicutt and D.Klatt 등의 'From Text to Speech' (Cambridge University Press, 1987) P.16 내지 P.150에 기재되어 있는 홀만트 합성 방식이나, 失頭의 「텍스트 음성 합성 기술의 최신 상황」(Interface, Dec., 1996) P.161 내지 P.165에 기재되어 있는 파형 중첩 방식을 쓸 수 있다. 물론, 다른 텍스트 음성 합성 방식을 사용할 수도 있다. 또한, 각 영역(英譯)에 대응하는 음성을 미리 녹음하여, 그대로 혹은 압축하여 메모리 등의 기억 장치에 저장하여 놓도록 해도 좋다.

계속해서, 음성 생성 수단(12)은, 메모리(15)상의 SYNWAVE 207에 저장된 역어 음성의 파형 데이터를 아날로그 데이터 혹은 패킷 데이터로 변환하고, 음성 출력 수단(13)을 통하여 음성으로서 전화 단말(1)에 보내어, 상기 전화 단말(1)이 갖는 메모리에 상기 보내어진 역어 음성을 저장한다. 다만, 음성 출력 수단(13)으로부터 출력된 역어 음성을 상기 전화 단말(1)이 갖는 메모리에 저장하지 않고, 자동 통역 서버(1000)의 메모리(15)상에 저장하도록 해도 좋다.

이 시점에서, 요금 부과 테이블(33)의 유저 ID 컬럼(411)의 상기 인증 서버(31)로부터 전송된 ID 셀에 대응하는 요금 컬럼 412의 내용에 통역 1회분의 소정의 요금이 가해진다. 예를 들면 통역 일회당 50엔이라는 요금을 설정하고 있으면, 전번의 예에서는, 유저 ID가 '1236'이므로, 요금 부과 테이블(33)의 유저 ID 컬럼411의 셀 중 '1236'인 셀(403)에 대응하는 요금 컬럼412의 셀이 '50'이 된다. 이 요금은, 1통역 서비스당 금액으로 설정해도 좋고, 어느 정액 요금으로 몇번이나 통역 서비스를 이용할 수 있도록 설정해도 좋다. 서비스 종료 후, 가산된 요금은, 각 ID의 유저에 대하여 청구되어 지불이 행하여진다. 요금 중의 수수료 등으로서 소정의 금액이 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)의 소유자에게 지불되고, 남은 금액이 자동 통역 서버(1000)의 소유자에게 지불되도록 한다. 다만, 이것들의 요금의 분배 방법이나 분배처는 일례이고, 상황에 따라서 설정할 수가 있다.

여기서, 유저가 전화 단말(1)상의 역어 출력의 기능을 할당한 소정의 버튼을 누르면, 상기 전화 단말(1)상의 메모리에 저장되어 있는 역어 음성을 판독하여, 상기 전화 단말(1) 상의 스피커(100)로부터 상기 역어 음성이 출력되도록 한다. 또한 유저가 전화 단말(1)상의 역어 출력의 기능이 할당된 소정의 버튼을 누르면, 그 때마다 상기 전화 단말(1)상의 메모리에 저장되어 있는 역어 음성을 판독하여, 전화 단말(1)상의 스피커(100)로부터 상기 역어 음성이 출력되도록 한다.

상기 역어 음성을 출력하기 위한 방법에는, 전화 단말(1)상의 역어 출력의 기능이 할당된 소정의 버튼을 누르는 방법에 한하지 않고, 소정의 단어 혹은 구 혹은 문장을 유저가 발성 입력하는 방법을 이용하여도 좋다. 즉, 유저는, 상기 역어 음성의 출력을 지시하는 취지의 소정의 단어 혹은 구 혹은 문장을 출력 지시 음성으로서 전화 단말(1)상의 마이크(104)로부터 발성 입력하고, 자동 통역 서버(1000)중의 음성 인식 수단(6)이 상기 음성을 인식하여, 출력 지시 음성이면, 상기 전화 단말(1)상의 메모리에 저장되어 있는 역어 음성을 판독하고, 전화 단말(1)상의 스피커(100)로부터 상기 역어 음성이 출력되도록 한다. 다만, 음성 출력 수단(13)으로부터 출력된 역어 음성이 상기 전화 단말(1)이 갖는 메모리에 저장되어 있지 않은 경우에는, 자동 통역 서버(1000)의 메모리(15)상에 저장되어 있는 역어 음성을 상기 전화 단말(1)에 보내도록 한다.

또, 여기서 전술한 상기 전화 단말(1)상의 메모리에 저장되어 있는 역어 음성을 판독하여 전화 단말(1)상의 스피커(100)로부터 상기 역어 음성을 출력시키는 경우에는, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에 어떤 정보를 보내는 일은 없고, 따라서 요금 부과 서버(32)에 의해서 상기 유저에 요금 부과가 이루어지는 일은 없다. 이상의 동작의 흐름을 도 18, 도 19에 흐름도로 도시한다.

계속해서, 본 발명의 제2 실시예에 관해서 도면을 이용하여 설명한다. 도 17은 본 발명의 제2 실시예를

도시하는 자동 통역 서비스 시스템의 구성도이다. 제1 실시예와 같이, 본 실시예에 있어서의 서버는, 퍼스널 컴퓨터, 워크스테이션 등, 중앙 연산 장치, 메모리를 구비하고, 동 도면과 같은 구성을 할 수 있는 것이면 어느 것이나 좋고, 플랫폼의 종류가 본 발명의 적용 범위를 한정하는 것은 아니다. 또한, 본 실시예는 영어를 일본어로 통역하는 자동 통역 서비스이지만, 이것은 일례이고, 일영, 중일, 일중 등, 언어의 종류는 한정되지 않는다. 동 도면에 있어서, 참조부호 1은 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말, 2는 모바일 인터넷 액세스 서비스 패킷 망, 3은 모바일 인터넷 액세스 서비스 게이트웨이 서버, 4는 전화 회선 보드, 5는 음성 입력 수단, 6은 음성 인식 수단, 7은 언어 번역 수단, 8은 단어 사전, 9는 문법 테이블, 10은 언어 변환 테이블, 11은 음성 생성 수단, 12는 소편 세트, 13은 음성 출력 수단, 14는 중앙 연산 장치, 15는 메모리, 16은 언어 종류 표시 수단, 17은 장면 표시 수단, 18은 문례 표시 수단, 19는 인식 후보 표시 수단, 20은 문장 사전, 21은 커맨드 사전, 22는 언어 종류 테이블, 23은 장면 테이블, 24는 문례 테이블, 31은 인증 서버, 32는 요금 부과 서버, 33는 요금 부과 테이블, 34는 통화용 통신망, 1000은 자동 통역 서버, 1001은 모바일 인터넷 액세스 서비스 패킷 망에 접속되는 회선, 1002는 통화용 통신망에 접속되는 회선이다.

메모리(15)의 데이터 구조 및 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)의 개관의 예는 제1 실시예와 마찬가지로이다. 다만, 여기서는, 제1 실시예와는 달리, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말로서, 통화 음성 및 데이터를 동일한 프로토콜로 취급하지 않는 전화 단말을 가정하여 설명한다.

도 17, 도 3에 있어서, 먼저, 모바일 인터넷 액세스 서비스 이용이 가능한 전화 단말(1)의 전원(104)을 넣고, 모바일 인터넷 접속용 버튼(102)을 누르면, 모바일 인터넷 액세스 서비스 패킷망(2)을 통해, 센터의 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)에 접속되어, 인증 서버(31)로 모바일 인터넷 액세스 서비스 이용 등록 유저인 것이 확인된다. 이하, 장면 확정까지의 동작은, 제1 실시예와 마찬가지로이다.

장면이 확정되면, 다음에, 자동 통역 서버(1000)의 문례 표시 수단(18)이 기동되고, 문례 테이블(24)을 이용하여 회선(1001)을 통해 전화 단말(1)상의 디스플레이(101)에, 도 7과 같이 통역 가능한 문례를 표시함과 동시에, 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된다. 계속해서, 음성 입력 수단(5)은, 시스템을 음성 입력 가능한 상태로 한다. 그 후, 유저는, 상기 문례를 보면서, 예를 들면 전번의 예이면, 일본어로 레스토랑의 장면에서의 통역시키고 싶은 문장을 전화 단말(1)의 송화기의 마이크(104)를 향해서 발성 입력한다. 여기서는, 예로서, 레스토랑의 장면에서 「물을 원합니다」라는 문장을 영어로 통역하고 싶은 경우를 생각하는 것으로 한다. 문례 테이블(24)은 도 14와 같은 데이터 구조를 갖고 있고, 문례 표시 수단(18)은, 문례 테이블(24)의 장면 번호511의 각 항목 중, 메모리(15)상의 SCENE210에 저장되어 있는 값의 항목의 문례513를 문례 번호512의 '1'로부터 순차적으로, 상기 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 표시한다. 전번의 예에서는, 메모리(15)상의 SCENE210에 '3'이 저장되어 있기 때문에, 도 14의 문례 테이블(24)의 장면 번호511의 값이 '3'인 항목501, 502, 503, 504의 문례(513)의 내용 「안녕하십니까」, 「고맙습니다」, 「[ ]는 어디입니까」, 「[ ]를 원합니다」를 상기 전화 단말(1)에 보내어 도 7과 같이 상기 전화 단말(1)상의 디스플레이(101)에 M 개씩 순차 표시한다. M은 디스플레이(101)의 크기에 따라서 설정하는 상수이지만, 여기서는 M = 4로 한다.

유저는, 상기 문례의 중에 「[ ]를 원합니다」라는 문례가 있기 때문에, 이 문례를 기초로 「물을 원합니다」라고 발성 입력한다. 이 발성 입력에 즈음하여, 전화 단말(1)상의 발성 입력의 신호의 의미를 가 지게한 소정의 버튼을 눌러 자동 통역 서버(1000)의 음성 입력 수단(5)이 음성 입력 가능한 상태가 되고 나서 발성해도 좋고, 특히 버튼 등의 트리거를 이용하지 않고 자동 통역 서버(1000)의 음성 입력 수단(5)이 기동된 후 언제나 음성 입력 가능한 상태이더라도 좋다. 또, 표시하는 문례는, 상술한 예와 같이 슬롯[ ]을 함유한 문장 혹은 문법이라도 좋고, 혹은 문장 그 자체라도 좋다.

제2 실시예에서는, 통화 음성 및 데이터를 동일한 프로토콜로 다룰 수 없는 전화 단말을 가정하고 있기 때문에, 발성 입력한 음성을 자동 통역 서버에 보내기 위해서는, 통화용 통신망(20)을 통한 필요가 있다. 따라서, 유저는, 상기 전화 단말(1)로부터, 모바일 인터넷 액세스 서비스 게이트웨이 서버(3)를 경유하고 회선(1001)을 통하여 자동 통역 서버(1000)에 접속한 것과는 다른 전화번호로 통화용 통신망(34)을 경유하고 회선(1002)을 통해 자동 통역 서버에 접속한다. 여기서, 유저가 다시 접속하는 방법 외에, 자동 통역 서버(1000)의 음성 입력 수단(5)이 상기 유저의 전화 단말(1)에 접속하도록 해도 좋다. 하여간, 유저가 발성한 음성은, 통화용 통신망(34)을 경유하고 회선(1002)을 통해 자동 통역 서버(1000)에 보내어진다.

이하, 음성 인식 수단(6)이 기동되어 상기 발성 입력된 음성이 인식되고, 결과가 메모리에 저장되는 동작까지는, 제1 실시예와 마찬가지로이다. 다음에, 인식 후보 표시 수단(19)이 기동되어, RECOGS (1), RECOGS (2), ..., RECOGS (M)205의 내용을, 도 8과 같이, 회선(1001)을 통해 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차적으로 표시한다. 이 때, 메모리(15)상의 ICNT204에 '1'를 저장하고, 전화 단말(1)상의 디스플레이(101)에 표시되어 있는 RECOGS (ICNT)의 내용이 흑백 반전되도록 한다. M은 디스플레이(101)의 크기에 따라서 설정하는 상수이지만, 여기서는 M = 4로 한다. 또한, 메모리(15)상의 INDEX215에 '0'을 저장한다.

유저는, 제1 후보의 표시가 자신이 발성한 내용과 일치하고 있는지 혹은 가까우면, 그대로 확정 기능이 할당된 소정의 버튼을 눌러 확정한다. 제1 후보의 표시가 정확하지 않은 경우, 유저는, 상하 커서 이동 기능이 할당된 소정의 버튼을 누름으로써, 전화 단말(1) 상의 디스플레이(101)상의 정확한 문자열이 표시되어 있는 부분에까지 흑백 반전 표시 영역을 아래쪽으로 이동시킨다. 즉, 하 이동 버튼을 누를 때마다, 메모리(15)상의 ICNT 204의 값을 인크리먼트하여, 회선(1001)을 통해 전화 단말(1)의 디스플레이(101)상의 동일하게 메모리(15)상의 RECOG(ICNT)의 내용이 표시되어 있는 부분만을 흑백 반전 표시하도록 한다. 여기서, ICNT 204의 값이 M을 넘은 경우에는, 메모리(15)상의 INDEX215의 값에 'M'을 가산하고, 다음 M 개의 후보 RECOGS (INDEX+1), RECOGS (INDEX+2), ..., RECOGS (INDEX+M)를 메모리(15)로부터 판독하여, 회선(1001)을 통해 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차

적으로 표시한다. 이때, 메모리(15)상의 ICNT204에 '1'을 저장하고, 전화 단말(1)상의 디스플레이(101)에 표시되어 있는 RECOGS(INDEX+1), RECOGS(INDEX+2), ..., RECOGS(INDEX+M)중의 ICNT 번째의 표시가 흑백 반전되도록 한다. 이하, 마찬가지로 하여, 다음의 M 개의 후보를, 전화 단말(1)에 보내어 상기 전화 단말(1)상의 디스플레이(101)에 순차적으로 표시하여 간다. 또, 도 15에 도시한 음성 인식에 사용하는 문장 사전(20)의 구조는 일례이고, 예를 들면, 문법과 단어 사전을 조합하는 방법도 있다. 그 경우도, 메모리(15)상의 SCENE210에 저장된 장면 번호에 대응시켜, 장면별 문법이나 사전을 이용하도록 한다.

자신이 발성한 내용의 확인 동작은, 제1 실시예와 마찬가지로이지만, 자동 통역 서버(1000) 중의 음성 생성 수단(12)은, 전화 단말(1)상의 다음 후보 제시 기능이 할당된 소정의 버튼이 눌리면, 메모리(15)상의 JCNT208를 인크리먼트하여, RECOGS(JCNT)를 판독하고, 상기 문자열을 합성 음성으로 변환하며, 상기 음성의 파형 데이터를 디지털/아날로그 변환에 의해 아날로그 데이터로 변환하여, 음성 출력 수단(13)을 통해 음성으로서 회선(1001)을 통해 전화 단말(1)에 보낸다.

확정 버튼을 누르면, 메모리(15)상의 RECOGS (ICNT)에 저장되어 있는 문자열을 동일하게 메모리(15)상의 JAPANESE203에 저장한다. 확정 버튼의 신호를 자동 통역 서버(1000)에 보내는 루트는, 회선(1001)을 경유해도 좋고, 회선(1002)을 경유해도 좋다. 이 확정 버튼을 누르는 대신에, 어느 소정의 단어 혹은 구 혹은 문장을 유저가 발성 입력하는 방법을 이용하여도 좋다. 즉, 유저는, 상기 전송된 음성을 전화 단말(1)상의 스피커 입력(100)으로부터 듣고, 상기 음성이 자신이 발성 입력한 내용과 합치하고 있으면, 합치하고 있다는 취지의 소정의 단어 혹은 구 혹은 문장을 확정 음성으로서 전화 단말(1)상의 마이크(104)로부터 발성 입력하고, 상기 음성이 자신이 발성 입력한 내용과 합치하지 않으면, 합치하고 있는 경우와는 다른 소정의 단어 혹은 구 혹은 문장을 비확정 음성으로하여 전화 단말(1)상의 마이크(104)로부터 발성 입력하도록 하며, 회선(1002)을 통하여 자동 통역 서버(1000)에 보내어지고, 상기 자동 통역 서버 중의 음성 인식 수단(6)이, 상술한 바와 같은 문장 입력과 마찬가지로 음성을 인식하여, 전체 매칭 정도가 미리 설정한 임계치를 하회하는 경우 혹은 ICNT204의 값이 N을 넘은 경우, 커맨드 사전(21)과의 대조를 개시한다. 다음에, 자동 통역 서버(1000) 중의 언어 번역 수단(7)이 기동되지만, 상기 언어 번역 수단(7)의 번역 동작은, 제1 실시예와 마찬가지로이다.

다음에, 언어 번역 수단(7)은, 도 10에 도시한 바와 같이, 메모리(15)의 JAPANESE203 및 RESULT206에 각각 저장되어 있는 내용을 회선(1001)을 통하여 전화 단말(1)상의 디스플레이(101)에 표시한다. 동 도면의 표시 방법은 일례이고, 표시의 방법이 본 발명의 범위를 한정하는 것은 아니다.

계속해서, 자동 통역 서버(1000) 중의 음성 생성 수단(12)이 기동되지만, 상기 음성 생성 수단(12)이 음성을 생성하는 동작은 제1 실시예와 마찬가지로이다.

계속해서, 음성 생성 수단(12)은, 메모리(15)상의 SYNWAVE207에 저장된 역어 음성의 파형 데이터를 아날로그 데이터 혹은 패킷 음성으로 변환하고, 음성 출력 수단(13)을 통해 음성으로서 회선(1002)을 통해 전화 단말(1)에 보내며, 상기 전화 단말(1)이 갖는 메모리에 상기 전송된 역어 음성을 저장한다. 이하, 요금 부과의 방법 및 동작에 관해서도, 제1 실시예와 마찬가지로이다.

### 발명의 효과

본 발명에 따르면, 휴대 전화 단말을 이용한 정확도가 높은 자동 통역 서비스를 실현할 수 있기 때문에, 상대의 언어를 의식하지 않고 이국인끼리의 상호 커뮤니케이션을 용이하게 수행할 수 있다.

### (57) 청구의 범위

#### 청구항 1

음성 통역 서비스 방법에 있어서,

통역 서버와 모바일 인터넷 액세스망으로 접속된 휴대 단말의 화면에 상기 통역 서버에 저장된 통역용 문례집(通譯用文例集)을 표시시키는 단계와,

상기 휴대 단말로부터 음성 통역 서비스로의 접속 지시를 받아 상기 휴대 단말과 상기 통역 서버와의 접속을 전화망 접속으로 전환하는 단계와,

상기 휴대 단말에 입력된 유저 음성을 전화망을 통하여 수신하는 단계와,

상기 유저 음성을 인식하는 단계와,

인식한 유저 음성을 다른 언어 표현으로 통역하는 단계와,

통역된 다른 언어 표현에 대응하는 음성을 전화망을 통하여 상기 휴대 단말에 출력하는 단계를 포함하는 것을 특징으로 하는 음성 통역 서비스 방법.

#### 청구항 2

제1항에 있어서,

상기 모바일 인터넷 액세스망을 통하여, 상기 유저가 통역 가능한 언어 종류를, 상기 휴대 단말의 표시 화면에 메뉴로서 표시시키는 단계와,

유저로부터 상기 언어 종류 메뉴로부터 통역 희망 언어의 선택을 받는 단계와,

인식한 유저 음성을 선택된 언어의 표현으로 통역하는 단계를

포함하는 것을 특징으로 하는 음성 통역 서비스 방법.



### 청구항 3

제1항에 있어서,

상기 모바일 인터넷 액세스망을 통하여 통역 가능한 문례 분야를 상기 휴대 단말의 표시 화면에 표시시키는 단계와,

유저로부터 상기 분야 메뉴로부터 통역 가능 문례 장면의 선택을 받는 단계와,

인식한 유저 음성을 선택된 문례 장면에 기초한 언어 표현으로 통역하는 단계를

포함하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 4

제1항에 있어서, 유저 음성을 인식한 결과를 전화망을 통하여 유저의 휴대 단말에 음성 출력하는 단계를 포함하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 5

제4항에 있어서,

상기 인식한 결과를 휴대 단말로 수신한 유저로부터의 지시를 받은 후, 통역 음성을 전화망을 통하여 상기 휴대 단말에 출력하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 6

제5항에 있어서, 상기 지시는, 상기 휴대 단말의 버튼을 밀어 내림에 따라 행해지고, 상기 유저가 상기 버튼을 누를때마다 통역 음성을 출력하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 7

제5항에 있어서, 상기 지시는, 상기 유저에 의한 특정한 단어 혹은 문장 음성의 발성 입력에 의해 행해지고, 상기 지시에 의해 통역 음성을 출력하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 8

제7항에 있어서, 상기 유저에 의한 특정 단어 혹은 문장 음성을 발성하고, 입력을 받을 때 마다, 통역 음성을 출력하는 것을 특징으로 하는 통역 서비스 방법.

### 청구항 9

제1항에 있어서,

음성 인식에 이용하는 사전 혹은 문법 혹은 이들 양쪽을 상기 휴대 단말의 표시 화면에 표시된 문례에 관한 항목에 조립하여 음성 인식을 행하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 10

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 통역 음성을 상기 휴대 단말에 출력하는 횟수에 무관하게, 유저에게 요금을 부과하는 것을 특징으로 하는 음성 통역 서비스 방법.

### 청구항 11

음성 통역 서버에 있어서,

전화망을 통하여 음성의 입력을 받는 음성 입력 수단과,

입력된 음성을 인식하여 소정의 기호열로 변환하는 음성 인식 수단과,

변환된 기호 열을 상기 입력된 음성의 언어와 다른 언어의 표현으로 변환하는 언어 변환 수단과,

변환된 언어 표현에 대응하는 음성을 출력하는 음성 출력 수단과,

통역용 문례를 저장하는 문례 테이블과,

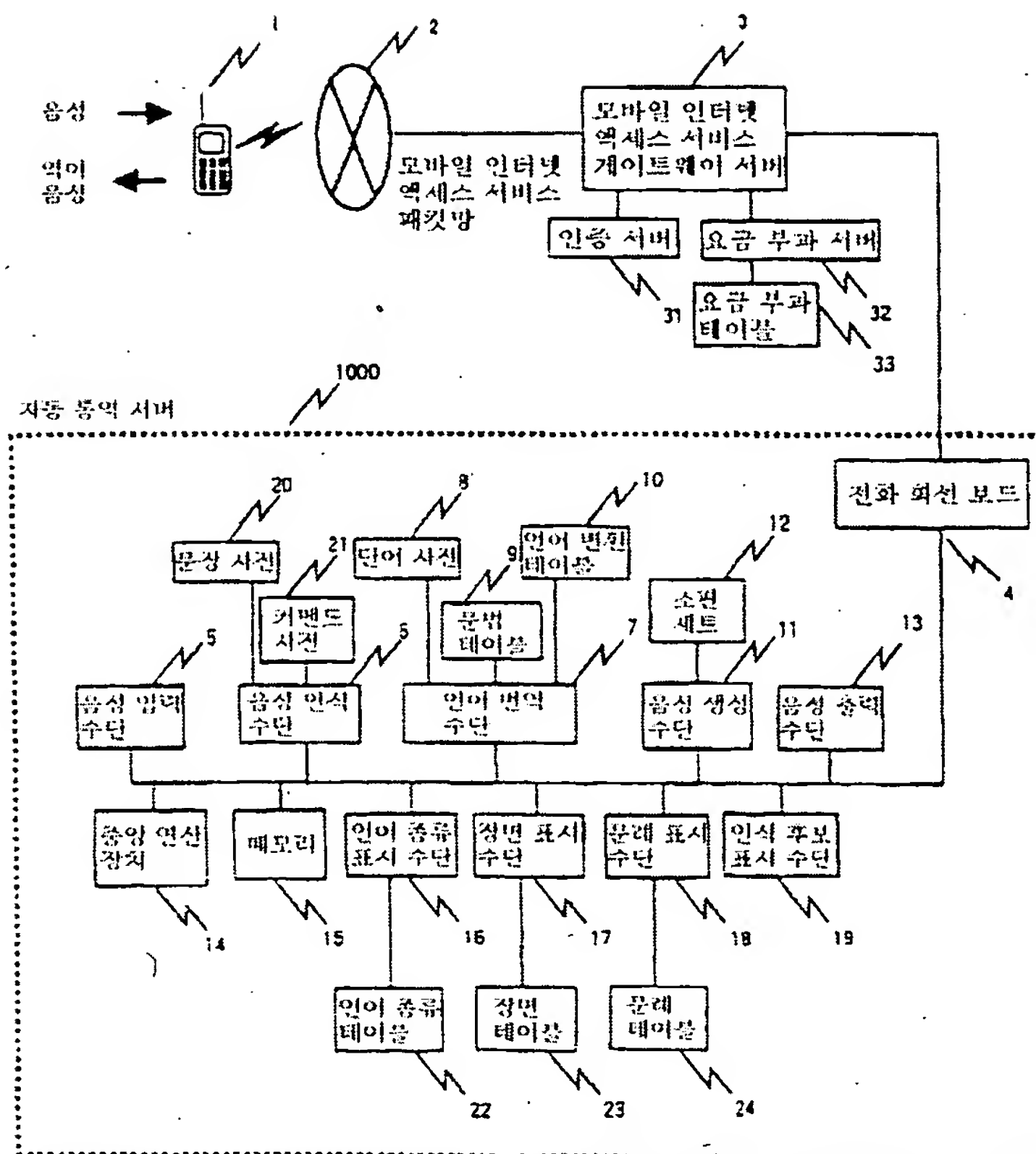
모바일 인터넷 액세스망으로 접속된 휴대 단말의 화면에 상기 통역용 문례를 표시시키는 문례 표시 수단을 포함하는 것을 특징으로 하는 음성 통역 서버.

### 청구항 12

제11항에 있어서, 상기 음성 인식 수단은, 상기 문례 테이블을 참조하여 통역용 문례와, 동작을 지시하는 커맨드문을 구별하여 인식하는 것을 특징으로 하는 음성 통역 서버.

도면

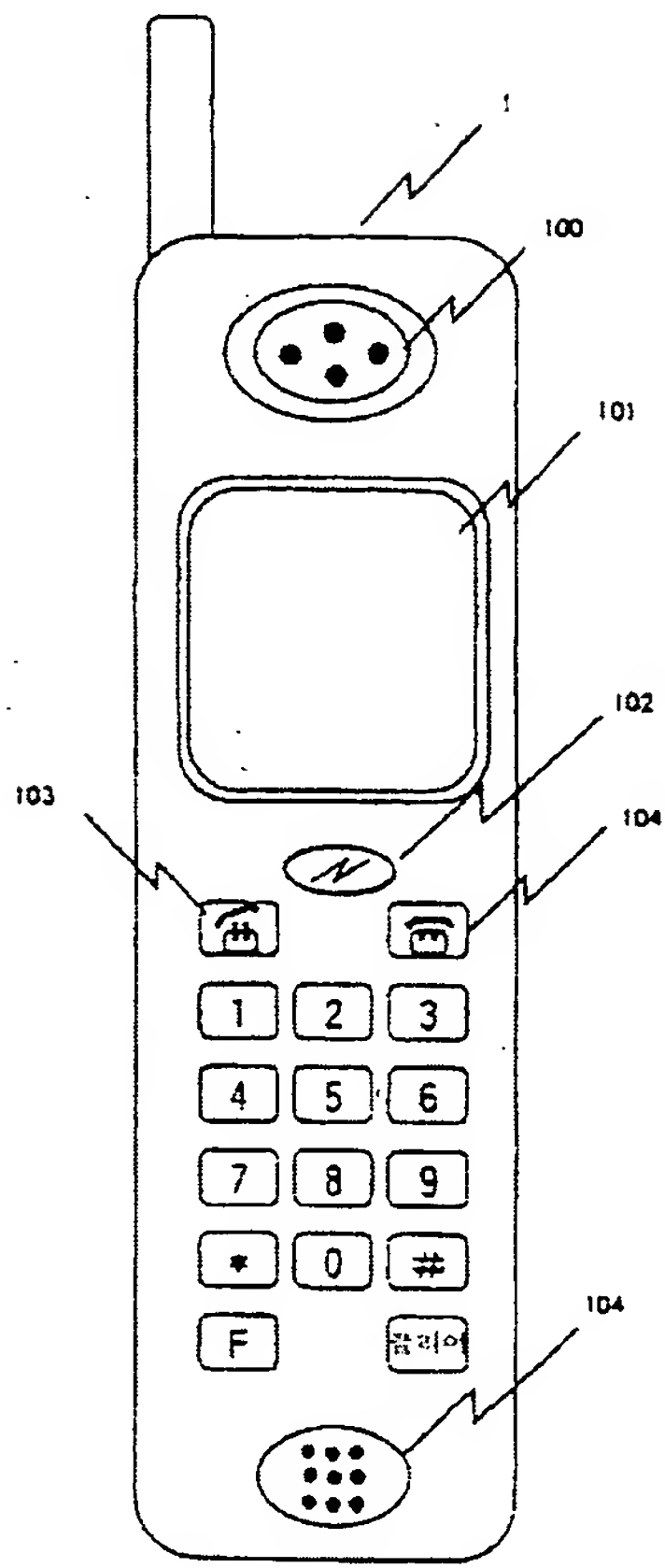
도면 1



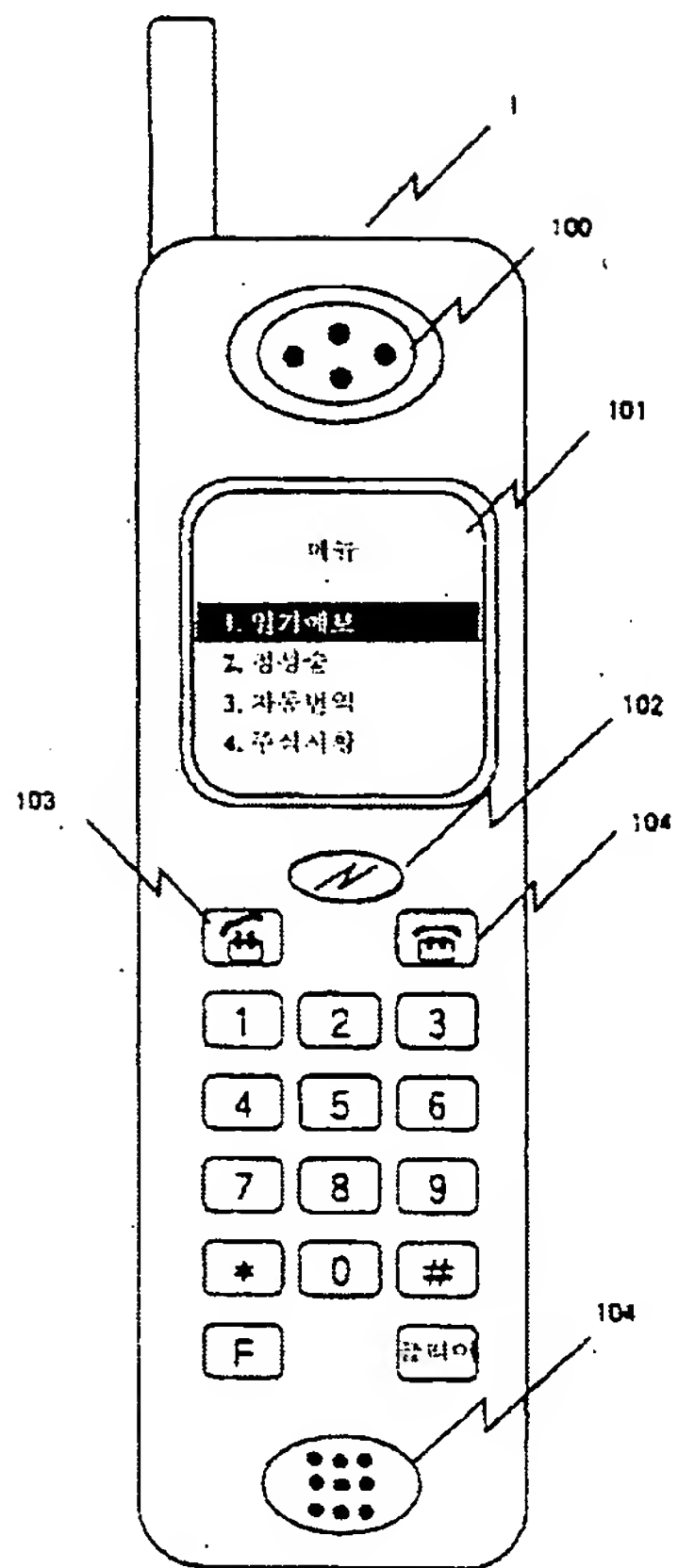
522

201	WAVE	
202	CSTRING	
203	JAPANESE	
204	ICNT	
205	RECOGS(1)	
	RECOGSNUM(1)	
	RECOGSNUM(1)	
	:	
	RECOGS(N)	
	RECOGSNUM(N)	
	RECOGSNUM(N)	
206	RESULT	
207	SYNWAVE	
208	JOINT	
209	LANG	
210	SCENE	
211	BSSENTENCE	
212	CSSENTENCE	
213	PCAND	
214	SCAND	
215	INDEX	
	:	

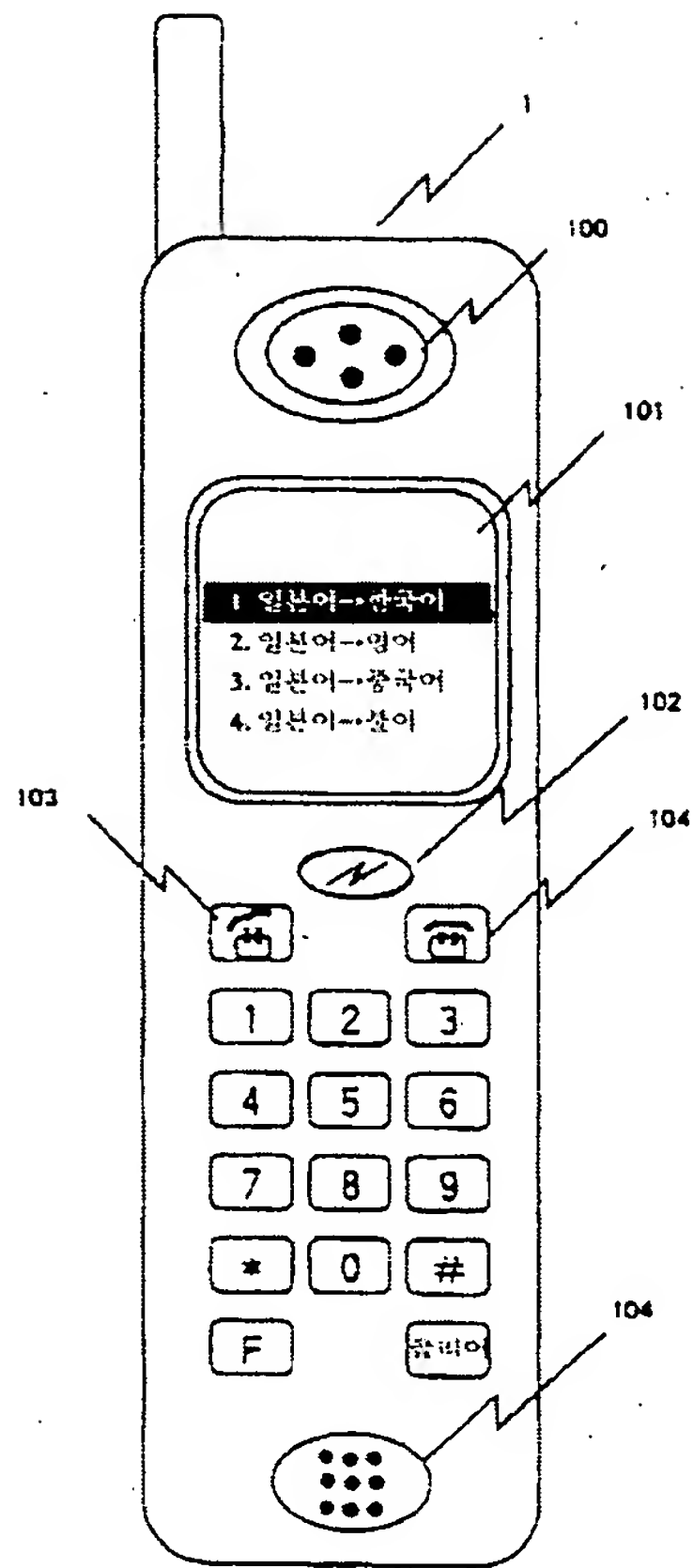
도면3



도면4

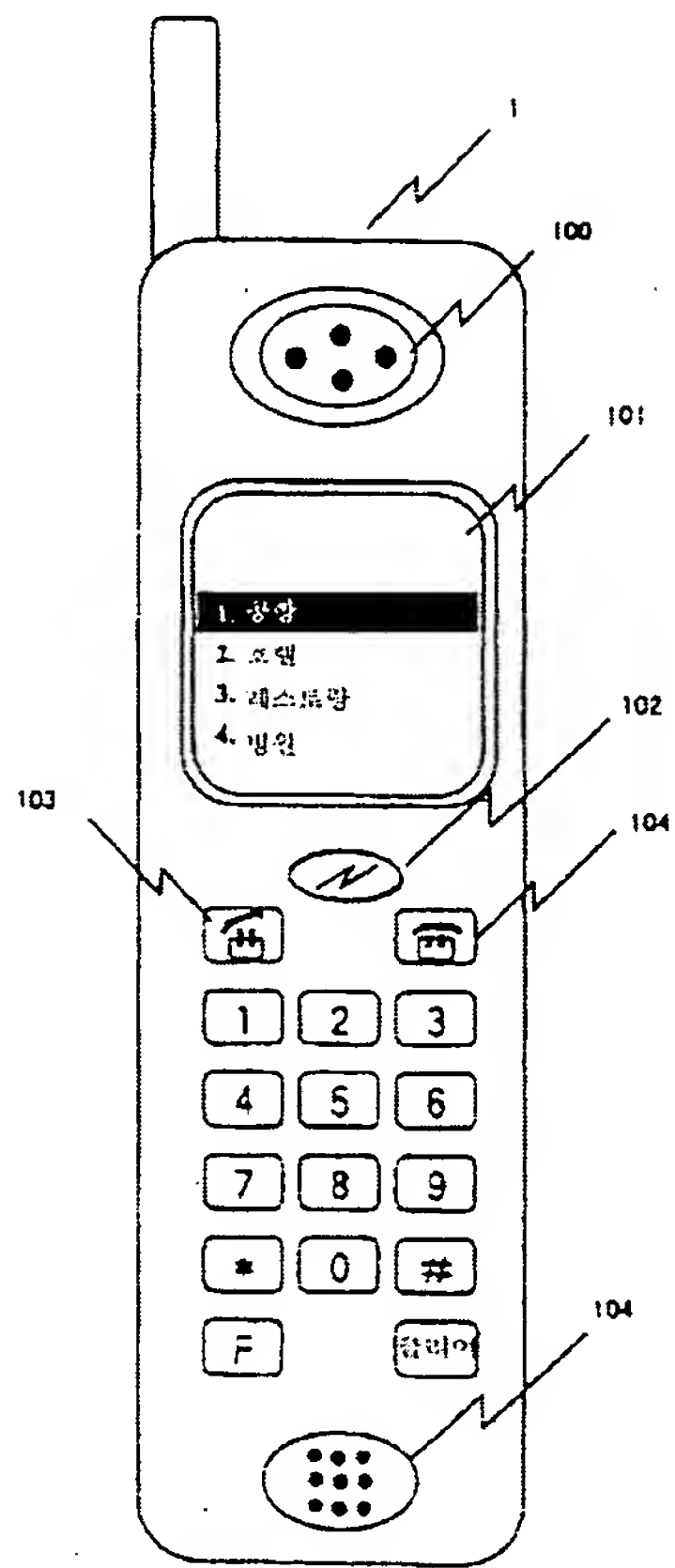


도면5

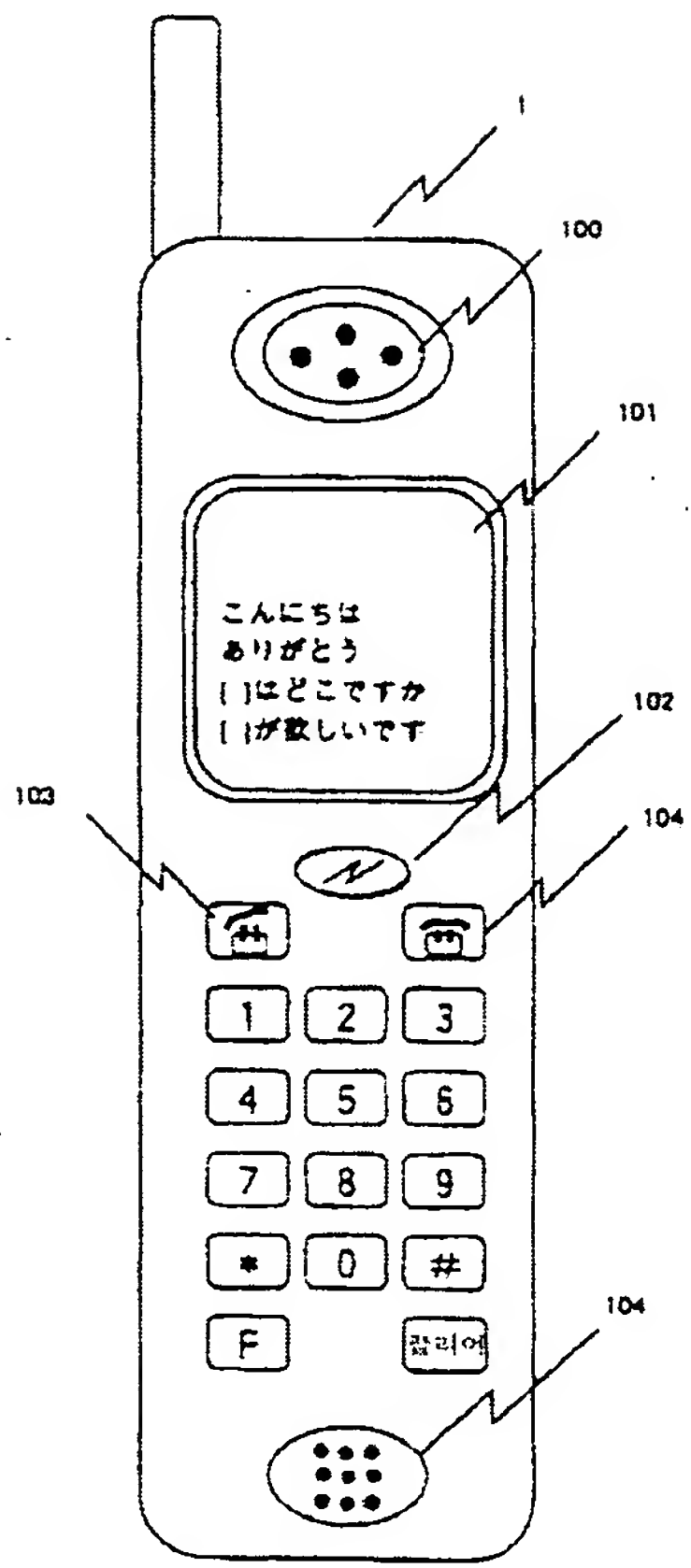




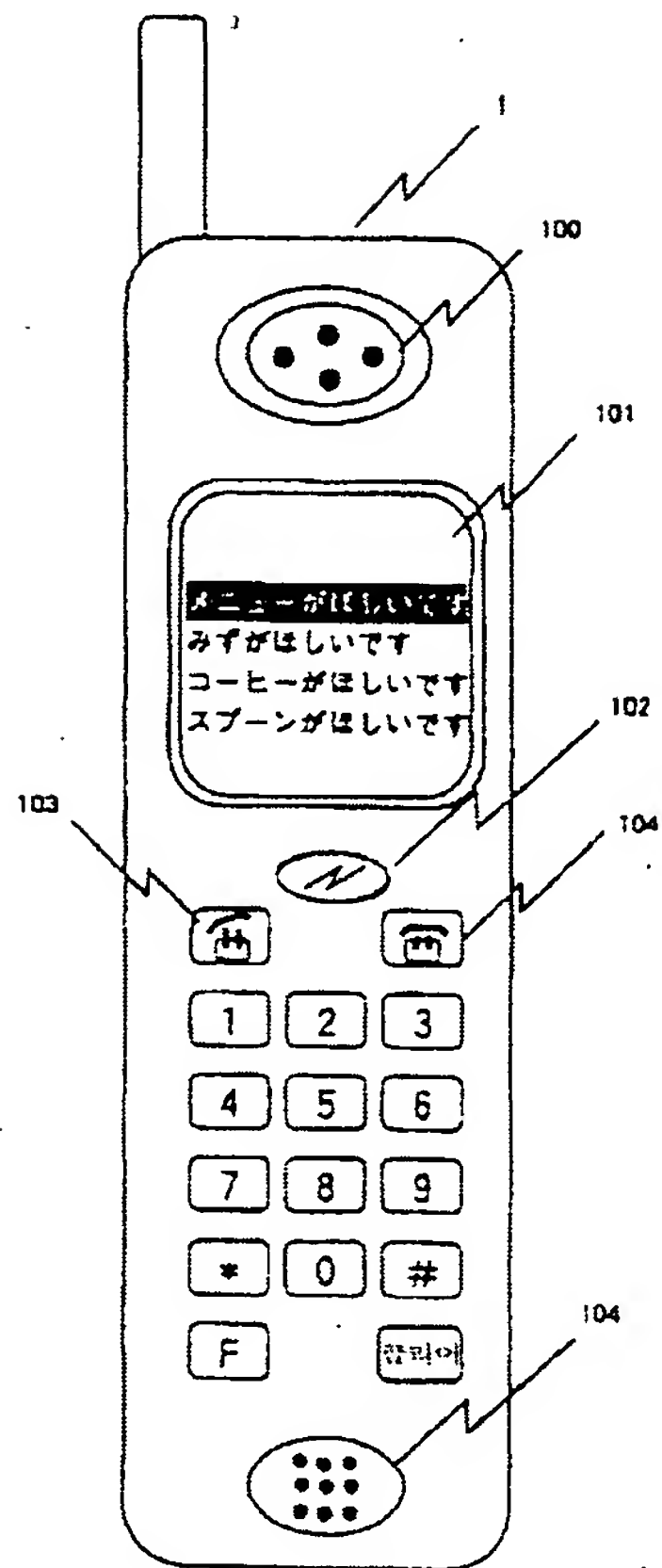
도면6



도면7



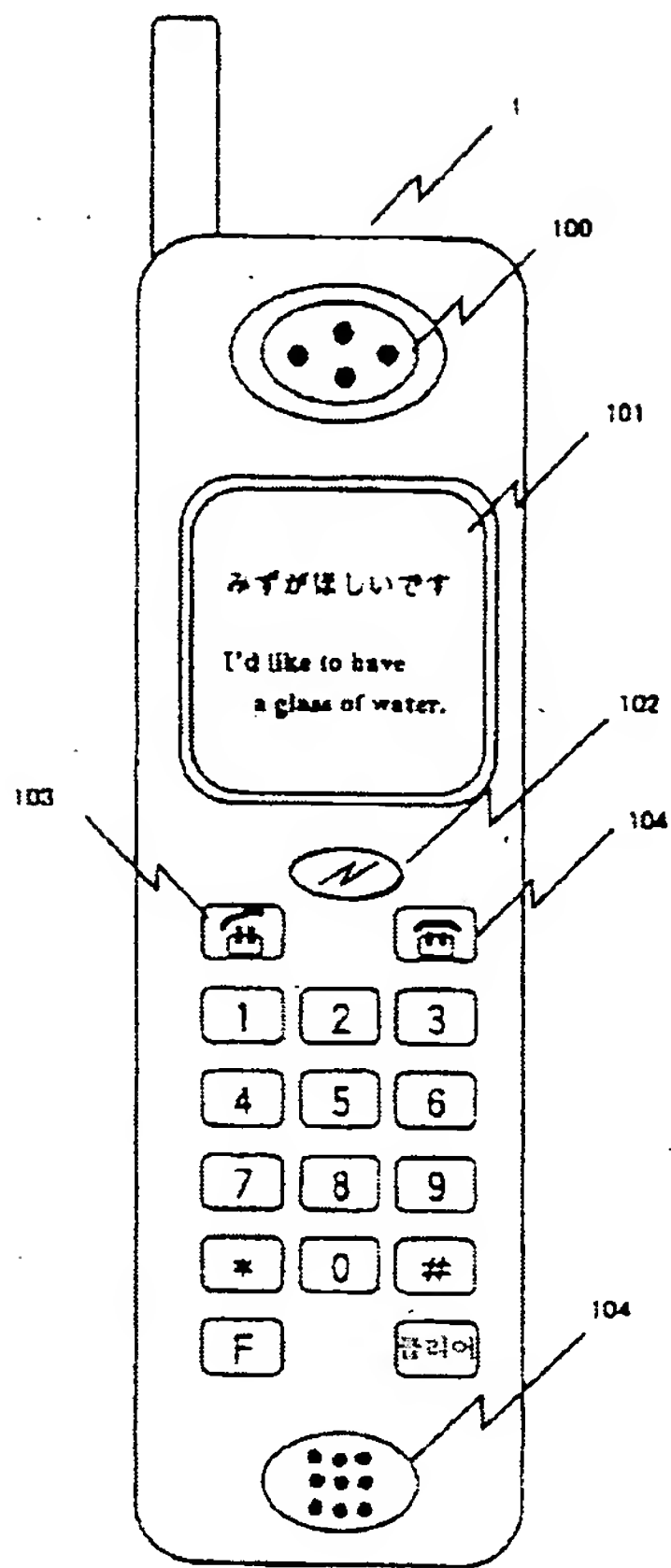
도면8



도면9

문례 번호	문장 번호	역어		
		1	2	...
:	:	:	:	:
4	1	...	I'd like to have a glass of water.	...
4	2	...	Can I see the menu?	...
4	3	...	I'd like to have a cup of coffee.	...
4	4	...	I'd like to have a spoon.	...
:	:	:	:	:

도면10



도면11

	411	412
	사용자 ID	부과 요금
401	:	:
402	1234	100
403	1235	0
404	1236	300
	1237	850
	:	:

도면12

	언어 종류 번호	언어 종류
801	1	일본어→한국어
802	2	일본어→영어
803	3	일본어→중국어
804	4	일본어→말어
	:	

도면13

	장면 번호	장면
901	1	공항
902	2	호텔
903	3	레스토랑
904	4	병원

도면14

	상면 번호	문래 번호	문래
	1	1	[ ]はどこですか
	:	:	
501	3	1	こんにちは
502	3	2	ありがとう
503	3	3	[ ]はどこですか
504	3	4	[ ]が欲しいです
505	3	5	[ ]はありますか
506	3	6	[ ]てください
	:	:	

도면 15

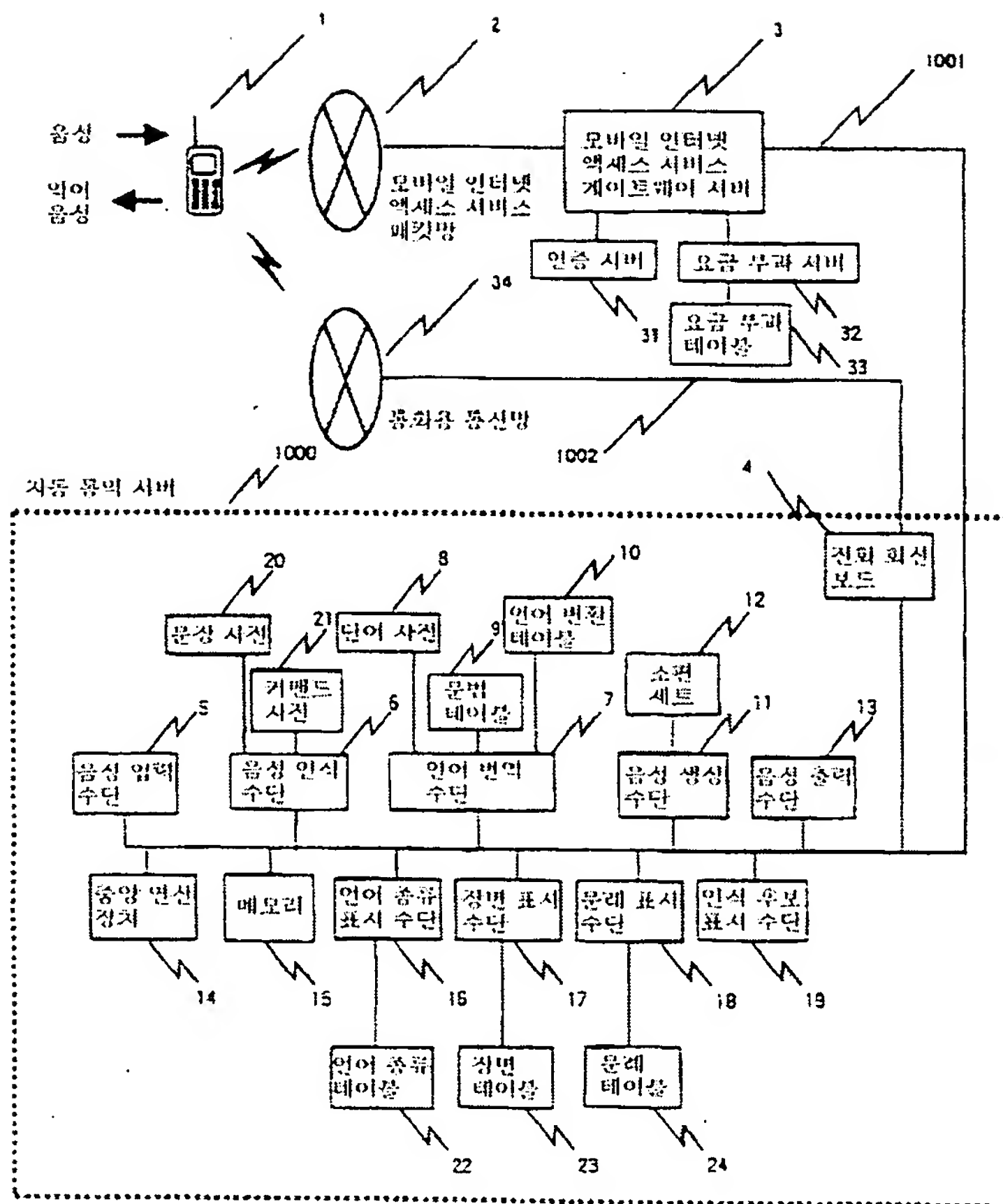
	문제 번호	문장 번호	문장	특정 벡터 시계열
601	1	1	こんにちは	A11, A12, ...
602	2	1	ありがとう	A21, A22, ...
603	:	:	:	:
604	4	1	みずがほしいです	A31, A32, ...
605	4	2	のにゅーがほしいです	A41, A42, ...
606	4	3	こーひーがほしいです	A51, A52, ...
	4	4	すぶーんがほしいです	A61, A62, ...
	:	:	:	:

도면 16

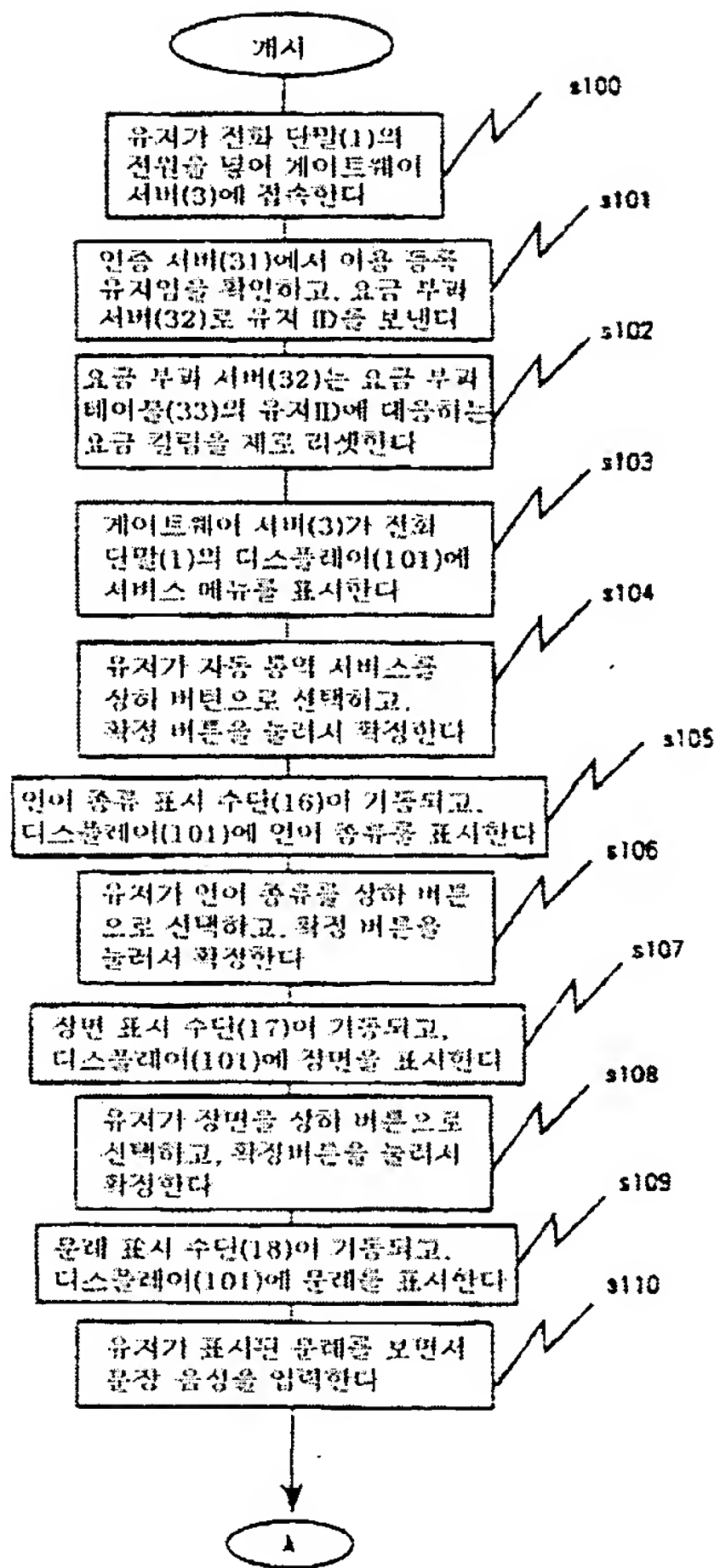
	커맨드 번호	커맨드	특정 벡터 시계열
701	1	かくてい	B11, B12, ...
702	2	つぎ	B21, B22, ...
703	3	サービス終了	:
	:		B31, B32, ...



도연17



도면 18



도면 19

